

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 04:29:55 : Search time 2346.74 Seconds  
(without alignments)  
17786.465 Million cell updates/sec

Title: US-10-030-390-3

Perfect score: 8241

Sequence: 1 agatctaacatccaagaacg.....aaaataacagttattatcg 8241

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	6561.6	79.6	9276	10	US-09-776-191-45
2	6561.6	79.6	9276	15	US-10-190-030B-13
3	6561.6	79.6	9276	15	US-10-302-840A-13
4	6561.6	79.6	9276	15	US-10-287-219-13
5	6561.6	79.6	9276	15	US-10-112-221A-11
6	6561.6	79.6	9276	15	US-10-104-271-11
7	6561.6	79.6	9276	16	US-10-147-211A-13
8	6383.8	77.5	9776	10	US-09-883-848A-31
9	5709.8	69.3	10462	10	US-09-883-848A-34
10	5665.8	68.8	10491	10	US-09-883-848A-32
11	5657.8	68.7	10512	10	US-09-883-848A-32
12	5142.8	62.4	8584	15	US-10-038-722-71
13	5130.8	62.3	8590	15	US-10-038-722-75
14	4565.2	55.4	8157	15	US-10-038-722-70
					Sequence 45, Appl
					Sequence 13, Appl
					Sequence 13, Appl
					Sequence 13, Appl
					Sequence 11, Appl
					Sequence 11, Appl
					Sequence 13, Appl
					Sequence 31, Appl
					Sequence 34, Appl
					Sequence 32, Appl
					Sequence 71, Appl
					Sequence 75, Appl
					Sequence 70, Appl

Query Match 79.6%; Score 6561.6; DB 10; Length 9276;

Best Local Similarity 84.7%; Pred. No. 0;

Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

#### ALIGNMENTS

#### RESULT 1

US-09-776-191-45  
; Sequence 45, Application US/09776191  
; Publication No. US20030119168A1  
; GENERAL INFORMATION:  
; APPLICANT: Edwin L. Madison  
; APPLICANT: Edgar O. Ong  
; APPLICANT: Giann-Chern Yeh  
; APPLICANT: Corvas International, Inc.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: TRANSMEMBRANE SERINE PROTEASES, THE ENCODED PROTEINS AND  
; TITLE OF INVENTION: METHODS BASED THEREON  
; FILE REFERENCE: 24745-1607  
; CURRENT APPLICATION NUMBER: US/09/776,191  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/213,124  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/234,840  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/179,982  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 60/183,542  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: 09/657,968  
; PRIOR FILING DATE: 2000-02-08  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 9276  
; TYPE: DNA  
; ORGANISM: Pichia pastoris

C	15	2659	32.3	2659	15	US-10-179-046-19	Sequence 19, Appl
C	16	2404.6	29.2	5740	10	US-09-921-380-3	Sequence 3, Appl
C	17	2320.8	28.2	5934	17	US-10-612-410-2	Sequence 2, Appl
C	18	2320.8	28.2	7618	17	US-10-612-410-1	Sequence 1, Appl
C	19	2314	28.1	11918	17	US-10-612-224-2	Sequence 2, Appl
C	20	2314	28.1	13227	17	US-10-612-224-4	Sequence 4, Appl
C	21	2314	28.1	13278	17	US-10-612-224-3	Sequence 3, Appl
C	22	2312.4	28.1	11466	17	US-10-466-959-4	Sequence 4, Appl
C	23	2302	27.9	4899	17	US-10-600-790-19	Sequence 19, Appl
C	24	2300.4	27.9	6865	15	US-10-338-411-33	Sequence 33, Appl
C	25	2300.4	27.9	6865	16	US-10-389-640-33	Sequence 33, Appl
C	26	2300.4	27.9	7894	15	US-10-338-411-34	Sequence 34, Appl
C	27	2300.4	27.9	7894	16	US-10-389-640-34	Sequence 34, Appl
C	28	2299.6	27.9	11152	9	US-09-847-101B-12	Sequence 12, Appl
C	29	2299.6	27.9	11152	10	US-09-482-682-12	Sequence 12, Appl
C	30	2298.8	27.9	8068	15	US-10-153-006-27	Sequence 27, Appl
C	31	2298.8	27.9	8068	15	US-10-159-006-35	Sequence 35, Appl
C	32	2298.2	27.9	7617	13	US-10-401-000-1	Sequence 1, Appl
C	33	2298.2	27.9	7626	13	US-10-401-000-2	Sequence 2, Appl
C	34	2298.2	27.9	9511	9	US-09-897-006-34	Sequence 34, Appl
C	35	2298.2	27.9	9511	10	US-09-897-511A-34	Sequence 34, Appl
C	36	2298.2	27.9	9511	13	US-10-397-079-34	Sequence 34, Appl
C	37	2297.8	27.9	5865	14	US-10-098-035-3	Sequence 3, Appl
C	38	2297.4	27.9	8687	17	US-10-467-546-15	Sequence 15, Appl
C	39	2297.4	27.9	8687	17	US-10-467-546-16	Sequence 16, Appl
C	40	2297.4	27.9	9362	17	US-10-467-546-18	Sequence 18, Appl
C	41	2297.4	27.9	9400	17	US-10-467-546-17	Sequence 17, Appl
C	42	2297.2	27.9	11641	13	US-10-684-300-13	Sequence 13, Appl
C	43	2297.2	27.9	11641	13	US-09-993-059-33	Sequence 33, Appl
C	44	2297.2	27.9	11641	15	US-10-103-327-33	Sequence 33, Appl
C	45	2297.2	27.9	11641	17	US-10-684-349-13	Sequence 13, Appl

QY	1	AGATCTAA	CACTCAAA	AGACGAA	AGGTTGA	TGAATGA	AACTCTTTT	GGCCATCCG	ACATCCAC	AG 60	
DB	1	AGATCTAA	CATCCAA	AGACGAA	AGGTTGA	TGAATGA	AACTCTTTT	GGCCATCCG	ACATCCAC	AG 60	
QY	61	GTCCATTCT	CACACAT	TAAGTCC	AAACGCA	CAGGAGG	GGATACAC	TACATAGC	ACACGAC	CGT 120	
DB	61	GTCCATTCT	CACACAT	TAAGTCC	AAACGCA	CAGGAGG	GGATACAC	TACATAGC	ACACGAC	CGT 120	
QY	121	TGCAAA	CGCAGG	ACCTCC	ACTCTCT	CTCTCT	CAACAC	CCACCTTTT	GGCCATCG	AAAAACC 180	
DB	121	TGCAAA	CGCAGG	ACCTCC	ACTCTCT	CTCTCT	CAACAC	CCACCTTTT	GGCCATCG	AAAAACC 180	
QY	181	AGCCAG	TATTTGG	CGTTGAT	TGGAGT	CTCGCTCAT	TCCAA	TTCCTTCTAT	TAGGCTACT	A 240	
DB	181	AGCCAG	TATTTGG	CGTTGAT	TGGAGT	CTCGCTCAT	TCCAA	TTCCTTCTAT	TAGGCTACT	A 240	
QY	241	ACACCAT	GACTTTATT	AGCTGTCT	ATCTCT	GGCCCCCTGG	CCGAGG	CTTCATG	TTTGTTTA	300	
DB	241	ACACCAT	GACTTTATT	AGCTGTCT	ATCTCT	GGCCCCCTGG	CCGAGG	CTTCATG	TTTGTTTA	300	
QY	301	TTTCCG	AAATGCA	ACAGCTCCG	ATTAC	ACCCGAA	CATCAT	CTCCAGAT	GAGGCTTT	CG 360	
DB	301	TTTCCG	AAATGCA	ACAGCTCCG	ATTAC	ACCCGAA	CATCAT	CTCCAGAT	GAGGCTTT	CG 360	
QY	361	AGTGTG	GGGTCAA	ATAGTTT	CATGTTT	CCCCAAAT	GGCCCCAA	ATCGAC	AGTTTAA	ACGCT 420	
DB	361	AGTGTG	GGGTCAA	ATAGTTT	CATGTTT	CCCCAAAT	GGCCCCAA	ATCGAC	AGTTTAA	ACGCT 420	
QY	421	GTCTTGA	AACTAA	TATGAC	AAAAAG	CGTGATCT	CATCC	AAAGATGA	ACTAAG	TTGTTCCG 480	
DB	421	GTCTTGA	AACTAA	TATGAC	AAAAAG	CGTGATCT	CATCC	AAAGATGA	ACTAAG	TTGTTCCG 480	
QY	481	TTGAAAT	GTCTAA	CGCCAG	TTGTTCA	AAAAAGAA	ACTTCC	AAAAAGT	CGCCAT	ACCGTTGT 540	
DB	481	TTGAAAT	GTCTAA	CGCCAG	TTGTTCA	AAAAAGAA	ACTTCC	AAAAAGT	CGCCAT	ACCGTTGT 540	
QY	541	CTTCTT	TGGTATT	GATTGAC	GAATGCT	CAAAAA	TAAATCT	CATTAA	TGCTT	PAGCGCAGTCT 600	
DB	541	CTTCTT	TGGTATT	GATTGAC	GAATGCT	CAAAAA	TAAATCT	CATTAA	TGCTT	PAGCGCAGTCT 600	
QY	601	CTCTAT	CGCTTCT	GAA	CCCCGGT	GCACCTGT	CGGAA	ACGCAAT	GGGAAA	CACCCGCT 660	
DB	601	CTCTAT	CGCTTCT	GAA	CCCCGGT	GCACCTGT	CGGAA	ACGCAAT	GGGAAA	CACCCGCT 660	
QY	661	TTTTGG	ATGATTAT	GCAAT	TGTTGCT	CCACAT	TGTATGCTT	CCAA	GAATCT	GGTGGGAATCT 720	
DB	661	TTTTGG	ATGATTAT	GCAAT	TGTTGCT	CCACAT	TGTATGCTT	CCAA	GAATCT	GGTGGGAATCT 720	
QY	721	GCTGAT	PAGCCTAA	CGTTCA	TGATCA	AAATTTA	CTGTTCT	TAA	CCCCCTAC	TGTGACAGCAAT 780	
DB	721	GCTGAT	PAGCCTAA	CGTTCA	TGATCA	AAATTTA	CTGTTCT	TAA	CCCCCTAC	TGTGACAGCAAT 780	
QY	781	ATAATA	ACAGAA	GAGCTG	CGCTGTCT	TAAACCTTTTT	TTTTAT	CATCAT	TATTAG	CTT 840	
DB	781	ATAATA	ACAGAA	GAGCTG	CGCTGTCT	TAAACCTTTTT	TTTTAT	CATCAT	TATTAG	CTT 840	
QY	841	ACTTTC	ATAAT	TGGC	ATGTTCCAA	TGACAG	CTTTGAT	TTTAA	CGACTTTT	TAACGA 900	
DB	841	ACTTTC	ATAAT	TGGC	ATGTTCCAA	TGACAG	CTTTGAT	TTTAA	CGACTTTT	TAACGA 900	
QY	901	CAACT	TGAGA	GAGATCA	AAAAAC	AACTAA	TATTCGA	AGGATCC	AAACGAT	GAGATTTCC 960	
DB	901	CAACT	TGAGA	GAGATCA	AAAAAC	AACTAA	TATTCGA	AGGATCC	AAACGAT	GAGATTTCC 960	
QY	961	TCAA	TTTTT	TACTG	CAGTTT	TATTCG	CAGATCC	CTCCG	GAT	TAGTGCTCCAG	TCAACACT 1020
DB	961	TCAA	TTTTT	TACTG	CAGTTT	TATTCG	CAGATCC	CTCCG	GAT	TAGTGCTCCAG	TCAACACT 1020
QY	1021	ACAC	CAGA	AGATGA	AAACG	GCACAA	ATTC	CGGCTGA	AGCTG	TCATCGG	TACTCAGATT 1080
DB	1021	ACAC	CAGA	AGATGA	AAACG	GCACAA	ATTC	CGGCTGA	AGCTG	TCATCGG	TACTCAGATT 1080
QY	1081	GAAG	GGGATTT	CGATG	TGTGCTTT	TG	CCATTTT	CCAA	CAGCA	CAAA	TAAACGGGTTAT 1140

	D	b	1081	G A A G G G G A T T C G A T G T T G C T G T T T T G C C A T T T C C A A C A G C A C A A T T A A C G G G T A T T G	1141
	Q	y	1141	T T T A T A A A T A C T A C T A T T G C C A G C A T T G C T T A A A G A A G A A G G G T A T C T C T C G A G A A A	1200
	D	b	1141	T T T A T A A A T A C T A C T A T T G C C A G C A T T G C T C T A A A G A A G A A G G G T A T C T C T C G A G A A A	1200
	Q	y	1201	A G A G A G C T G A A G C C A G G C C C A G G C C C A G G C C C A G G C C C A G A A A C A T G A T A T C A T G	1260
	D	b	1201	A G A G A G C T G A A G C C T A-----	1217
	Q	y	1261	G C C C C C G G A G A G A T A A A T T G T G C T T C C C C G G T G T C A C C C C C A G C A G T G C A C G A G	1320
	D	b	1218	-----	1217
	Q	y	1321	A G A G T T G C T G T T T T G A T G A C A G T G T C C G G G A T T C C C G T G T G C T T C C A C C C A T G C C	1380
	D	b	1218	-----	1217
	Q	y	1381	A T C G A G A A C A C T C A A G N A G A A A T G T C C C T C T A A C T A G T G G C G T A G A A T T C C C T A G G G	1440
	D	b	1218	-----C G T A G A A T T C C C T A G G G	1234
	Q	y	1441	C G G C C G C G A A T T A A T T C G C C T T A G A C A T A C T G T T C C T C A G T T C A A G T T G G G C A C T T A C G	1500
	D	b	1235	C G G C C G C G A A T T A A T T C G C C T T A G A C A T A C T G T T C C T C A G T T C A A G T T G G G C A C T T A C G	1294
	Q	y	1501	A G A A G A C C G G T C W T G T A G A T T C T A A T C A G A G A T G T C A G A T G C C A T T T G C T G A G A G	1560
	D	b	1295	A G A A G A C C G G T C T T G T A G A T T C T A A T C A G A G A T G T C A A A T G C C A T T T G C T G A G A G	1354
	Q	y	1561	A T C A G S C T T C A T T T T T G A T A C T T T T T A T T T G T A A C C T A T A G T A T A G A T T T T T T T T	1620
	D	b	1355	A T C A G S C T T C A T T T T T G A T A C T T T T T A T T T G T A A C C T A T A G T A T A G A T T T T T T T T	1414
	Q	y	1621	G T C A T T T G T T C T T C T C G T A C A G A C T T G C T C T G A T C A G C C T A T C T C G A G C T G A T A A	1680
	D	b	1415	G T C A T T T G T T C T T C T C G T A C A G A A T T A A G T G A G A A G T T C G T T T G T C A A G C T T A T C G A T A A	1474
	Q	y	1681	T A T C T T G T G T A G G G T T C G G A A A T C A I T T C A G A T T T G A T G T T T T C T T G G T A T T T C C C	1740
	D	b	1475	T A I T T T G T G T A G G G T T T G G A A A T C A T T C A G A T T T G A T G T T T T C T T G G T A T T T C C C	1534
	Q	y	1741	A C T C C T C T T C A G A G T A C A G A A G T T A A G T C A G A A G T T C G T T T G T C A A G C T T A T C G A T A A	1800
	D	b	1535	A C T C C T C T T C A G A G T A C A G A A G T T A A G T C A G A A G T T C G T T T G T C A A G C T T A T C G A T A A	1594
	Q	y	1801	G C T T T A A T C G G T A G T T A T C A G A T T A A A T T G C T A A C G A G T C A G A C C C G T G T A T G A A	1860
	D	b	1595	G C T T T A A T C G G T A G T T A T C A G A T T A A A T T G T A A C G A G T C A G A C C G T G T A T G A A	1654
	Q	y	1861	A T C T A A A A T G C G C T C A T C G T C A T C C T C G G A C C G T C A C C C T G G A T G C T P A G C A T A G G	1920
	D	b	1655	A T C T A A C A A T G C G C T A T C G T C A T C C T C G G A C C G T C A C C C T G G A T G C T P A G C A T A G G	1714
	Q	y	1921	C T T G G T T A T G C C G T A C T G C C G G C C T T T G C G G G A T A T G T C C A I T T C C G A C A G A T C G C	1980
	D	b	1715	C T T G G T T A T G C C G T A C T G C C G G C C T T T G C G G G A T A T G T C C A I T T C C G A C A G A T C G C	1774
	Q	y	1981	C A G T C A C T A T T G C G T G C T P A G G C T A T A T G C G T T C A T G C A A T T T C T A T G C G A C C C G T	2040
	D	b	1775	C A G T C A C T A T T G C G T G C T P A G G C T A T A T G C G T T C A T G C A A T T T C T A T G C G A C C C G T	1834
	Q	y	2041	T C T C G A G A C A T G T C C A C C G C T T T G C C C G C C A G T C T G C T G C T T G C G T A C T T G G	2100
	D	b	1835	T C T C G A G A C A T G T C C A C C G C T T T G C C C G C C C A G T C T G C T G C G T T C G C T A C T T G G	1894
	Q	y	2101	A G C C A C T A T C G A C T A C C G A T C A T G G C G A C C A C A C C G T C G T G T G A T A T C A A A T C T A	2160
	D	b	1895	A G C C A C T A T C G A C T A C C G A T C A T G G C G A C C A C A C C G T C G T G T G A T A T A T C G A A T C T A	1954
	Q	y	2161	A A T G T A A G T T A A A A T C T A A A T A A T T A A A T A A G T C C A G T T T C T C A A T A G A A C C T T A A	2220





QY 4381 TAGCGGCAAGGATGCAACTTTATCAACACGAGGCTCTATAGATAACCCCTAGGCGCTGGGA 4440  
 Db 4175 TAGCGGCAAGGATGCAACTTTATCAACACGAGGCTCTATAGATAACCCCTAGGCGCTGGGA 4234  
 QY 4441 TCATCCTTTGGACACACTTTCTGCGCAATCTAGTCTCCCAAAATCACTTCAATGATACCAT 4500  
 Db 4235 TCATCCTTTGGACACACTTTCTGCGCAATCTAGTCTCCCAAAATCACTTCAATGATACCAT 4294  
 QY 4501 TATTGTACAACTTGAGCAAGTGTGCGATCAGCTCCCTCAAAATGCTCTCTGTAAACGATG 4560  
 Db 4295 TATTGTACAACTTGAGCAAGTGTGCGATCAGCTCCCTCAAAATGCTCTCTGTAAACGATG 4354  
 QY 4561 ACTCAACTTGGACACACTTTCTGCGCAATCTAGTCTCCCAAAATCACTTCAATGATACCAT 4620  
 Db 4355 ACTCAACTTGGACACACTTTCTGCGCAATCTAGTCTCCCAAAATCACTTCAATGATACCAT 4414  
 QY 4621 GCTGCTCAGCAGCATAGGAAACACGCGTTTCTACCAAACTCAAGGAATATCAAACT 4680  
 Db 4415 GCTGCTCAGCAGCATAGGAAACACGCGTTTCTACCAAACTCAAGGAATATCAAACT 4474  
 QY 4681 CTGCAACAATTGCGTATGCAAGTAGCAAGGAAATGTCACTTGAAGTCGACAGTGAG 4740  
 Db 4475 CTGCAACAATTGCGTATGCAAGTAGCAAGGAAATGTCACTTGAAGTCGACAGTGAG 4534  
 QY 4741 TGTAGTCTTGAGAAATCTGAGCGGTATTTTATATCATGAGTCAGTCAGTCATCAGGAGA 4800  
 Db 4535 TGTAGTCTTGAGAAATCTGAGCGGTATTTTATATCATGAGTCAGTCAGTCATCAGGAGA 4594  
 QY 4801 TCCTCTACGCGGACGATCGTGCCGA 4828  
 Db 4595 TCCTCTACGCGGACGATCGTGCCGA 4654  
 QY 4829 4828  
 Db 4655 CTCGTGAAGAGTGTGCTGACTCATACAGGCGCTGAATCGGCCCATATCAGCGCAGA 4714  
 QY 4829 4828  
 Db 4715 AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGATGAGTGGACCACTGGTGATTTGA 4774  
 QY 4829 4828  
 Db 4775 ACTTTTGCTTTGCCACGGAACGGTCTGCGTTGTCGGAAGATGCGTGATCTGATCTCTTCA 4834  
 QY 4829 4828  
 Db 4835 ACTCAGCAAAAGTTCGATTATTCAACAAGCCGCTCCCGTCAAGTCAGCGTAATGCT 4894  
 QY 4829 4828  
 Db 4895 CTGCCAGTGTACAACTTAACCAATCTGATTAGAAAACTCATCGAGCATCAATG 4954  
 QY 4829 4828  
 Db 4955 AAACGTCAATTTATCATATCAGGATATCAATACCATATTTTGAAGAAAGCGTTTCTG 5014  
 QY 4829 4828  
 Db 5015 TAATGAAGGAGAAACTCACCGAGGAGTTCATAGGATGCAAGATCTCTGTTATCGGTC 5074  
 QY 4829 4828  
 Db 5075 TCGGATCCGACTCGTCCAAATCAATACAACTTAATTTCCCTCGTCAAAATAAG 5134  
 QY 4829 4828  
 Db 5135 GTTATCAAGTGAGAAATCACCATGAGTCAGCTGAATCGGTGAGAAATGCAAAAGCTT 5194  
 QY 4829 4828  
 Db 5195 ATGCATTTCTTCCAGACTTGTTCACAGGCGAGCCATTAACGCTCGTCAATAAATCACT 5254  
 QY 4829 4828

Db 5255 CGCATCAACCAACCGTTATTTCATTCGTGATTGGCCCTGAGCGAGCAAAATACGGATC 5314  
 QY 4829 4828  
 Db 5315 GCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGAACAACCTGCCAG 5374  
 QY 4829 4828  
 Db 5375 CGCATCAACCAATATTTTCACTGAAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTT 5434  
 QY 4829 4828  
 Db 5435 CCCGGGATCGCAGTGGTGTAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTTGAT 5494  
 QY 4829 4828  
 Db 5495 GGTCCGAAGAGGCATAAATTCGGTCAGCCAGTTTGTAGTCTGACCATCTCATCTGTAAACATC 5554  
 QY 4829 4828  
 Db 5555 ATTGCAACGCTACCTTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGCTTCCCAT 5614  
 QY 4829 4828  
 Db 5615 CAATCGATAGATTGTCCGACCTGATTGCGGACATTTATCGAGCCCATTTATACCCATA 5674  
 QY 4829 4828  
 Db 5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5734  
 QY 4829 4828  
 Db 5735 ATGGCTCATAAACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCTAGA 5794  
 QY 4829 4828  
 Db 5795 TGATATATTTTATCTTGTGCAATGTATACATCAGAGATTTTGAGACACAACGTGGCTTC 5854  
 QY 4829 4828  
 Db 5855 CCCCCCCCCCTGCAAGTTCGCGCATCACCGCGCCACAGGTGCGGTGCTGGCGCTATAT 5914  
 QY 4880 CCGCGACATACCGATGCGGAAGATCGGGCTCGGCATCTTCGGGCTCATGAGCGCTTGT 4939  
 Db 5915 CCGCGACATACCGATGCGGAAGATCGGGCTCGGCATCTTCGGGCTCATGAGCGCTTGT 5974  
 QY 4940 CCGCGTGGGTATGTTGCGAGGCCCGCTGGCGGGGAGTGTGGGCGGCAATCTCTTGTGA 4999  
 Db 5975 CCGCGTGGGTATGTTGCGAGGCCCGCTGGCGGGGAGTGTGGGCGGCAATCTCTTGTGA 6034  
 QY 5000 TGCAACATCTCTTGGCGGGCGGCTGCTCAACGGCTCAACCTACTACTTGGGCTCTCTCT 5059  
 Db 6035 TGCAACATCTCTTGGCGGGCGGCTGCTCAACGGCTCAACCTACTACTTGGGCTCTCTCT 6094  
 QY 5060 AATGAGGAGTCCGATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAATGGT 5119  
 Db 6095 AATGAGGAGTCCGATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAATGGT 6154  
 QY 5120 GATACCCGATCTTTCAGTGTCTTGAAGTCTCTATCAGATTTGCCCACTAAGCAAC 5179  
 Db 6155 GATACCCGATCTTTCAGTGTCTTGAAGTCTCTATCAGATTTGCCCACTAAGCAAC 6214  
 QY 5180 CCGAGGAGAGATTTTCAGTGTAAATTTCTCTGACTTTTGGTCTCAGTAGACTCGAATG 5239  
 Db 6215 CCGAGGAGAGATTTTCAGTGTAAATTTCTCTGACTTTTGGTCTCAGTAGACTCGAATG 6274  
 QY 5240 TGAGACTATCTCGTTATGACAGCAAGAAATGTCTTCTTGGAGACAGTAAATGAGTCCC 5299  
 Db 6275 TGAGACTATCTCGTTATGACAGCAAGAAATGTCTTCTTGGAGACAGTAAATGAGTCCC 6334  
 QY 5300 ACCAATAAGAAATCTTGTATTATCAGGAACAACCTTCTTGTTCGAACTTTTTCGGTCCC 5359

Db 6335 ACCAATAAAGAAATCCTTTGTTATACGGAACAAACTTCTTGTGTTTCCAACTTTTTCGGTGCC 6394  
Qy 5360 TTGAACTATAAATAGAGTGGATATGTCGGGTAGGAATGGAGCGGCAAAATGCTTAAC 5419  
Db 6395 TTGAACTATAAATAGAGTGGATATGTCGGGTAGGAATGGAGCGGCAAAATGCTTAAC 6454  
Qy 5420 TTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGACTGATGCAACTTCAGTGAAC 5479  
Db 6455 TTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGACTGATGCAACTTCAGTGAAC 6514  
Qy 5480 GTTGCTATTTCCGTTCAAAACCAATCCGAATCAGAGAAATCAAGTTGTTGTTCTACTATT 5539  
Db 6515 GTTGCTATTTCCGTTCAAAACCAATCCGAATCAGAGAAATCAAGTTGTTGTTCTACTATT 6574  
Qy 5540 GATCAAGCCAGTGGGCTTTGAACTGCAATAGTGTGCTGCTGTTTGTAGGTCATCTT 5599  
Db 6575 GATCAAGCCAGTGGGCTTTGAACTGCAATAGTGTGCTGCTGTTTGTAGGTCATCTT 6634  
Qy 5600 TGTATGAATAAATCTAGTCTTTGATCTAAATAAATCTTGAACGAGCAAGGCGATAAATACC 5659  
Db 6635 TGTATGAATAAATCTAGTCTTTGATCTAAATAAATCTTGAACGAGCAAGGCGATAAATACC 6694  
Qy 5660 CAATCTAAACTCTTTTAAACGTTAAAGGCAAGTATGCTGCCTGTATTAAACCCC 5719  
Db 6695 CAATCTAAACTCTTTTAAACGTTAAAGGCAAGTATGCTGCCTGTATTAAACCCC 6754  
Qy 5720 AAATCAGCTCTAGTCTGATCTCATCAACTGAGGGGCACTATCTTGTGTTAGAGAAAT 5779  
Db 6755 AAATCAGCTCTAGTCTGATCTCATCAACTGAGGGGCACTATCTTGTGTTAGAGAAAT 6814  
Qy 5780 TTGGGAGATGCGATATCGAGAAAAGGTACGCTGATTTTAAAGCGTGAATTTTATCTCAA 5839  
Db 6815 TTGGGAGATGCGATATCGAGAAAAGGTACGCTGATTTTAAAGCGTGAATTTTATCTCAA 6874  
Qy 5840 GATCTCTCCCTCGCGGCTTTTCGGTGATGACGGTGAAACCTCTGACACATGAGCTCCG 5899  
Db 6875 GATCTCTCCCTCGCGGCTTTTCGGTGATGACGGTGAAACCTCTGACACATGAGCTCCG 6934  
Qy 5900 GAGACGFTACAGCTTTGCTGTAGACGGATCCCGGAGCAGACAAAGCCGTCAGGGCGC 5959  
Db 6935 GAGACGFTACAGCTTTGCTGTAGACGGATCCCGGAGCAGACAAAGCCGTCAGGGCGC 6994  
Qy 5960 TCAGCGGTGTGGGGGTGTGGGGGCGACCCATGACCCAGTCAAGTACGATAGCGGA 6019  
Db 6995 TCAGCGGTGTGGGGGTGTGGGGGCGACCCATGACCCAGTCAAGTACGATAGCGGA 7054  
Qy 6020 GTGTATCTGCTTAACCTATCCGGCANTCAGACGAGATTGTACTGAGAGTGACCATATGC 6079  
Db 7055 GTGTATCTGCTTAACCTATCCGGCANTCAGACGAGATTGTACTGAGAGTGACCATATGC 7114  
Qy 6080 GGTGTGAATACCGCACAGATCGCTTAAGGAGAAATACCGCATCAGGCGCTCTTCCGCTT 6139  
Db 7115 GGTGTGAATACCGCACAGATCGCTTAAGGAGAAATACCGCATCAGGCGCTCTTCCGCTT 7174  
Qy 6140 CCTGCTCACTGACTCGTGGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6199  
Db 7175 CCTGCTCACTGACTCGTGGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7234  
Qy 6200 CAAAGGCGTAATACGGTTATCCACAGAAATCAGGGGATAGCGAGGAAAGAAATGTAG 6259  
Db 7235 CAAAGGCGTAATACGGTTATCCACAGAAATCAGGGGATAGCGAGGAAAGAAATGTAG 7294  
Qy 6260 CAAAGGCGTAATACGGTTATCCACAGAAATCAGGGGATAGCGAGGAAAGAAATGTAG 6319  
Db 7295 CAAAGGCGTAATACGGTTATCCACAGAAATCAGGGGATAGCGAGGAAAGAAATGTAG 7354  
Qy 6320 GGCTCCGCCCCCTGACAGCATCACAAATACGAGCTCAAGTCAGAGTGGCGAAACC 6379  
Db 7355 GGCTCCGCCCCCTGACAGCATCACAAATACGAGCTCAAGTCAGAGTGGCGAAACC 7414  
Qy 6380 CGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTGCTGGCTCTCCTG 6439  
Db 7415 CGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTGCTGGCTCTCCTG 7474

Qy 6440 TTCCGACCTGCGCTTAACGGATACCTGTCGGCTTTTCCCTTCGGGAAGCGTGCGC 6499  
Db 7475 TTCCGACCTGCGCTTAACGGATACCTGTCGGCTTTTCCCTTCGGGAAGCGTGCGC 7534  
Qy 6500 TTTCTCATAGCTCAAGCTGTAGTATCTCAGTTTCGGTGTAGTTCGTTCCGTCGAAGCTGG 6559  
Db 7535 TTTCTCAATGCTCAAGCTGTAGTATCTCAGTTTCGGTGTAGTTCGTTCCGTCGAAGCTGG 7594  
Qy 6560 GCTGTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCTGTC 6619  
Db 7595 GCTGTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCTGTC 7654  
Qy 6620 TTGAGTCCAAACCGGTAAACAGCACTTATCGCACTGCGAGAGCACTGGTAAACAGGA 6679  
Db 7655 TTGAGTCCAAACCGGTAAACAGCACTTATCGCACTGCGAGAGCACTGGTAAACAGGA 7714  
Qy 6680 TTAGCAGAGCAGGTATGTAGCGGTGTACAGAGTTCTTTGAAGTGTGCGCTTAACTAG 6739  
Db 7715 TTAGCAGAGCAGGTATGTAGCGGTGTACAGAGTTCTTTGAAGTGTGCGCTTAACTAG 7774  
Qy 6740 GCTACTAGAAAGGACAGTATTGTTGTTATCTGCTCTGCTGAAGCCAGTTACTTCGGAA 6799  
Db 7775 GCTACTAGAAAGGACAGTATTGTTGTTATCTGCTCTGCTGAAGCCAGTTACTTCGGAA 7834  
Qy 6800 AAGAGTTGGTAGCTCTTGCATCCGCAACAAACCAACCGCTGGTAGCGGTGTTTTTG 6859  
Db 7835 AAGAGTTGGTAGCTCTTGCATCCGCAACAAACCAACCGCTGGTAGCGGTGTTTTTG 7894  
Qy 6860 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAATCCTTTGATCTTT 6919  
Db 7895 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAATCCTTTGATCTTT 7954  
Qy 6920 CTACGGGCTGACCGCTCAGTGAGCAAGAACTCAGTTAAGGATTTTGGTCAAGAT 6979  
Db 7955 CTACGGGCTGACCGCTCAGTGAGCAAGAACTCAGTTAAGGATTTTGGTCAAGAT 8014  
Qy 6980 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATCT 7039  
Db 8015 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATCT 8074  
Qy 7040 AAGTATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAATCAGTAGGACCTTA 7099  
Db 8075 AAGTATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAATCAGTAGGACCTTA 8134  
Qy 7100 TCTCAGGATCTGCTATTTTCTGTTTCTCATCATAGTTGCTGCTGCTGCTGCTGCTGCTGCT 7159  
Db 8135 TCTCAGGATCTGCTATTTTCTGTTTCTCATCATAGTTGCTGCTGCTGCTGCTGCTGCTGCT 8194  
Qy 7160 CTACGATACGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCAGAGCCAC 7219  
Db 8195 CTACGATACGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCAGAGCCAC 8254  
Qy 7220 GCTCAGCGCTCAGATTTTATCAGCAATTAACAGCAGCGAGGCGGAGCGAGAA 7279  
Db 8255 GCTCAGCGCTCAGATTTTATCAGCAATTAACAGCAGCGAGGCGGAGCGAGAA 8314  
Qy 7280 GTGCTCCTCAACTTTATCCGCTCATCCAGTCTTAAATTTGTTGCGGGAGCTAGAG 7339  
Db 8315 GTGCTCCTCAACTTTATCCGCTCATCCAGTCTTAAATTTGTTGCGGGAGCTAGAG 8374  
Qy 7340 TTAGTAGTTTGGCAGTTAATAGTTTGGCAACGTTTGTGCTGCTGCTGCTGCTGCTGCTGCT 7399  
Db 8375 TTAGTAGTTTGGCAGTTAATAGTTTGGCAACGTTTGTGCTGCTGCTGCTGCTGCTGCTGCT 8434  
Qy 7400 TGTACGCTCGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 7459  
Db 8435 TGTACGCTCGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8494  
Qy 7460 TTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAGCTCTTCCGTCCTCCGATCGTTG 7519  
Db 8495 TTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAGCTCTTCCGTCCTCCGATCGTTG 8554

QY 7520 TCAGAGTAAAGTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATATTTCTC 7579  
Db 8555 TCAGAGTAAAGTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATATTTCTC 8614  
QY 7580 TTAAGTCAATCCGATCCGTAAGATGCTTTTCTGACCTGGTGGTCAACCAAGTCAT 7639  
Db 8615 TTAAGTCAATCCGTAAGATGCTTTTCTGACCTGGTGGTCAACCAAGTCAT 8674  
QY 7640 TCTGAGATAGTATGCGGACCGAGTGTCTTTTCCCGCGCTCAACACGCGATATA 7699  
Db 8675 TCTGAGATAGTATGCGGACCGAGTGTCTTTTCCCGCGCTCAACACGCGATATA 8734  
QY 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGAACTTCTCGGCGGAA 7759  
Db 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGAACTTCTCGGCGGAA 8794  
QY 7760 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCAGTACCCCACTCTGTCACCA 7819  
Db 8795 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCAGTACCCCACTCTGTCACCA 8854  
QY 7820 ACTGATCTTCAAGTCTTTTACTTTTCAACAGCTTTTCTGGGTGAGCAAAACAGGAAGC 7879  
Db 8855 ACTGATCTTCAAGTCTTTTACTTTTCAACAGCTTTTCTGGGTGAGCAAAACAGGAAGC 8914  
QY 7880 AAGATCCGCAAAAGGAAATAGGCGGACACCGGAATGTGATCTCATCTCTCTTCC 7939  
Db 8915 AAGATCCGCAAAAGGAAATAGGCGGACACCGGAATGTGATCTCATCTCTCTTCC 8974  
QY 7940 TTTTTCATATTAATGAAGCAATTTATCAGGCTTTTCTCATGCGGAGATACATATTTG 7999  
Db 8975 TTTTTCATATTAATGAAGCAATTTATCAGGCTTTTCTCATGCGGAGATACATATTTG 9034  
QY 8000 AATGATTTAGAAATTAAGAAATAGGCTTTTCCCGGACATTTCCCGGAAGTGGCCAC 8059  
Db 9035 AATGATTTAGAAATTAAGAAATAGGCTTTTCCCGGACATTTCCCGGAAGTGGCCAC 9094  
QY 8060 CTGAGCTCTAAGAAACCAATTTATCATGACATTAACCTATAAAATAGGCTATCACGA 8119  
Db 9095 CTGAGCTCTAAGAAACCAATTTATCATGACATTAACCTATAAAATAGGCTATCACGA 9154  
QY 8120 GGCCTTTGCTCTCAAGAAATTAATCTCATGTTTGCAGCTTATCATGATAGCTGAC 8179  
Db 9155 GGCCTTTGCTCTCAAGAAATTAATCTCATGTTTGCAGCTTATCATGATAGCTGAC 9214  
QY 8180 TCATGTTGTTATGTAATAGACGAGATCGGAAACACTGAAATAAAGAGTTATTT 8239  
Db 9215 TCATGTTGTTATGTAATAGACGAGATCGGAAACACTGAAATAAAGAGTTATTT 9274  
QY 8240 CG 8241  
Db 9275 CG 9276

## RESULT 2

US-10-190-030B-13  
; Sequence 13, Application US/10190030B  
; Publication No. US20030134298A1  
; GENERAL INFORMATION:  
; APPLICANT: Madison, Edwin  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING A  
; TITLE OF INVENTION: TRANSMEMBRANE SERINE PROTEASE 20, THE ENCODED POLYPEPTIDES AND  
; FILE OF INVENTION: METHODS BASED THEREON  
; FILE REFERENCE: 24745-1618  
; CURRENT APPLICATION NUMBER: US/10/190.030B  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 9276  
; TYPE: DNA  
; ORGANISM: Pichia pastoris  
US-10-190-030B-13

Query Match 79.6%; Score 6561.6; DB 15; Length 9276;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY 1 AGATCTAATCAATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAG 60  
Db 1 AGATCTAATCAATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAG 60  
QY 61 GTCCATTTCTACATAGTGCCTCACTCCCTCTCTCTCTCAACACCCACTTTTGGCATCGAAGAAC 120  
Db 61 GTCCATTTCTACATAGTGCCTCACTCCCTCTCTCTCTCAACACCCACTTTTGGCATCGAAGAAC 120  
QY 121 TGCAGACGAGACCTCCACTCCCTCTCTCTCTCAACACCCACTTTTGGCATCGAAGAAC 180  
Db 121 TGCAGACGAGACCTCCACTCCCTCTCTCTCTCAACACCCACTTTTGGCATCGAAGAAC 180  
QY 181 AGCCAGTTATTTGGCTTGAATGGAGCTCGCTCATTTCCCAATTCCTTCTATTAGGCTACTA 240  
Db 181 AGCCAGTTATTTGGCTTGAATGGAGCTCGCTCATTTCCCAATTCCTTCTATTAGGCTACTA 240  
QY 241 ACACCATGACTTTATAGCCTGTCTATCTCTGGCCCTCTGGCGAGGTTTCATGTTTCTTTA 300  
Db 241 ACACCATGACTTTATAGCCTGTCTATCTCTGGCCCTCTGGCGAGGTTTCATGTTTCTTTA 300  
QY 301 TTTCCGAATGCAACAGCTCCGCAATTTACCCGGAACATCATCTCCAGATGAGGGCTTTCTG 360  
Db 301 TTTCCGAATGCAACAGCTCCGCAATTTACCCGGAACATCATCTCCAGATGAGGGCTTTCTG 360  
QY 361 AGTGTGGGTCAAAATAGTTTTCATGTTTCCCAAAAGTCCCAAAAGTCCCAAAAGTCCCAAAAGT 420  
Db 361 AGTGTGGGTCAAAATAGTTTTCATGTTTCCCAAAAGTCCCAAAAGTCCCAAAAGTCCCAAAAGT 420  
QY 421 GTCTTGGAACTTAATATGACAAAGCGGTGATCTCTCAAGATGAACTAAGTTTGGTTG 480  
Db 421 GTCTTGGAACTTAATATGACAAAGCGGTGATCTCTCAAGATGAACTAAGTTTGGTTG 480  
QY 481 TTGAATAGTCTAAGCGGCGAGTGGTCAAAAGAAAGTTCCTCAAAAGTCCCAAAAGTCCCAAAAGT 540  
Db 481 TTGAATAGTCTAAGCGGCGAGTGGTCAAAAGAAAGTTCCTCAAAAGTCCCAAAAGTCCCAAAAGT 540  
QY 541 CTCTTTTGGTATGATGACAAATGCTCAAAATATCTCAATATGCTTATGCTTAGCGAGTCT 600  
Db 541 CTCTTTTGGTATGATGACAAATGCTCAAAATATCTCAATATGCTTATGCTTAGCGAGTCT 600  
QY 601 CTCTATCGCTTCTGAACCCCGGTGACCTGTGCGGAAAGCAAAATGGGGAACACCCGCT 660  
Db 601 CTCTATCGCTTCTGAACCCCGGTGACCTGTGCGGAAAGCAAAATGGGGAACACCCGCT 660  
QY 661 TTTTGGATGATATGATGCTCTCCCAATGATGCTTCAAGATTTCTGGTGGGAATACT 720  
Db 661 TTTTGGATGATATGATGCTCTCCCAATGATGCTTCAAGATTTCTGGTGGGAATACT 720  
QY 721 GCTGATAGCTTAAGCTTTCATGATCAAAATTTTAACTCTCTTAACTCTTAACTCTTAACTCTTAACT 780  
Db 721 GCTGATAGCTTAAGCTTTCATGATCAAAATTTTAACTCTCTTAACTCTTAACTCTTAACTCTTAACT 780  
QY 781 ATATAAACAAGAGAGAGCTGCTCTTAAACCTTTTATCATCATTTATGCTT 840  
Db 781 ATATAAACAAGAGAGAGCTGCTCTTAAACCTTTTATCATCATTTATGCTT 840  
QY 841 ACTTTTCAATATTCGAGTGTGCTTCAATTTGACAGCTTTTGAATTTTAAAGCTTTTAAAGCA 900  
Db 841 ACTTTTCAATATTCGAGTGTGCTTCAATTTGACAGCTTTTGAATTTTAAAGCTTTTAAAGCA 900  
QY 901 CAACTTGAGAGATCAAAACAACTAATTTTGAAGGATCCAAACGATGAGATTTCT 960  
Db 901 CAACTTGAGAGATCAAAACAACTAATTTTGAAGGATCCAAACGATGAGATTTCT 960  
QY 961 TCAATTTTACTGAGTTTTATTTGCGAGCATCTCTCGCATAGCTGCTCCAGTCAACT 1020  
Db 961 TCAATTTTACTGAGTTTTATTTGCGAGCATCTCTCGCATAGCTGCTCCAGTCAACT 1020

Qy	1021	ACAACAGAAGATGAAACGGCACAAATTC	CGCGTGAAGCTGTCATCGGTTACTCAGATT	TA 1080
Db	1021	ACAACAGAAGATGAAACGGCACAAATTC	CGCGTGAAGCTGTCATCGGTTACTCAGATT	TA 1080
Qy	1081	GAAGGGGATTTTCGATGTTGCTGTTTGG	CCATTTTCCAAACAGACAATAATACGGGTTAT	TG 1140
Db	1081	GAAGGGGATTTTCGATGTTGCTGTTTGG	CCATTTTCCAAACAGACAATAATACGGGTTAT	TG 1140
Qy	1141	TTTATAAATACTACTATTGCCAGATTTG	CTGTCTAAAGAAAGGGGTATCTCTCGAGAAA	1200
Db	1141	TTTATAAATACTACTATTGCCAGATTTG	CTGTCTAAAGAAAGGGGTATCTCTCGAGAAA	1200
Qy	1201	AGAGAGCTGAAGCCAGCGCCAGGCCAGG	CCAGGCCCGAGGAACAATGATCATG 1260	
Db	1201	AGAGAGCTGAAGCTTA-----	----- 1217	
Qy	1261	GCOCGCCGAGAGGATPAAAATTGGCGT	TCCCGGTGTACCGCCAGCAGTGCACGAG	1320
Db	1218	-----	----- 1217	
Qy	1321	AGAGTTGCTGTTTGTATGACAGTGTCCG	GGGATTCGCGTGGTGCTTCCACCCCATGG	CC 1380
Db	1218	-----	----- 1217	
Qy	1381	ATCGAGAACACTCAAGAAGAAGATGTCC	CTTAACTAGTGGCGTAGAATTCCTTAGGG	1440
Db	1218	-----	-----CGTAGAATTCCTTAGGG 1234	
Qy	1441	CGCGCCGAATTAATTCGCCCTTAGACATG	ACTGTTCTCAGTTCAAGTTGGGCACTTAG	C 1500
Db	1235	CGCGCCGAATTAATTCGCCCTTAGACATG	ACTGTTCTCAGTTCAAGTTGGGCACTTAG	C 1294
Qy	1501	AGAAGACCGGTCCTTGCTAGATTTCTAAT	CAAGAGATGT CAGAAATGCCATTTGGCTC	GAGAG 1560
Db	1295	AGAAGACCGGTCCTTGCTAGATTTCTAAT	CAAGAGATGT CAGAAATGCCATTTGGCTC	GAGAG 1354
Qy	1561	ATGAGGCTTCATTTTGATACATTTTTTAT	TATTGTAACTATAAGTATAGGATTTTTTTT	1620
Db	1355	ATGAGGCTTCATTTTGATACATTTTTTAT	TATTGTAACTATAAGTATAGGATTTTTTTT	1414
Qy	1621	GTCAATTTGTTCTTCCTGTAACAGCTTGC	TCTGATCAGCCCTATCTCGCAGCTGATGA	A 1680
Db	1415	GTCAATTTGTTCTTCCTGTAACAGCTTGC	TCTGATCAGCCCTATCTCGCAGCTGATGA	A 1474
Qy	1681	TATCTTGTGTGAGGGTTTCGGAAAAATCA	TTCCAGTTTGATGTTTTCTTGTTATTTCC	C 1740
Db	1475	TATCTTGTGTGAGGGTTTCGGAAAAATCA	TTCCAGTTTGATGTTTTCTTGTTATTTCC	C 1534
Qy	1741	ACTCCCTTCAGAGTACAGAGATTAAGTG	AGAGTTTAAAGTGTGTCAGCTTATCGATA	A 1800
Db	1535	ACTCCCTTCAGAGTACAGAGATTAAGTG	AGAGTTTAAAGTGTGTCAGCTTATCGATA	A 1594
Qy	1801	GCITTAATCGGTAGTTTATCACAGTTAA	ATTCCTAACCGCAGTCAAGCCCGTGTATGA	A 1860
Db	1595	GCITTAATCGGTAGTTTATCACAGTTAA	ATTCCTAACCGCAGTCAAGCCCGTGTATGA	A 1654
Qy	1861	ATCTACAATGCGCTCATCGTTCATCTCG	GCACCGTCACTTCGATGTAGGCATAGG	1920
Db	1655	ATCTACAATGCGCTCATCGTTCATCTCG	GCACCGTCACTTCGATGTAGGCATAGG	1714
Qy	1921	CTTGGTTATGCCGCTACTGCGGGCTCTT	TGCGGGATATCGTCCATTCGACAGCATCC	G 1980
Db	1715	CTTGGTTATGCCGCTACTGCGGGCTCTT	TGCGGGATATCGTCCATTCGACAGCATCC	G 1774
Qy	1981	CAGTCACTATGGCTGCTGCTAGCGCTAT	ATATGGTTGATGCAATTTCTATGCGCAC	CCCGT 2040
Db	1775	CAGTCACTATGGCTGCTGCTAGCGCTAT	ATATGGTTGATGCAATTTCTATGCGCAC	CCCGT 1834
Qy	2041	TCTCGGAGCACTGTCCAGCCGCTTTGG	CGCGCCCGCCAGTCTCTCGTCTGCTACT	TGG 2100
Db	1835	TCTCGGAGCACTGTCCAGCCGCTTTGG	CGCGCCCGCCAGTCTCTCGTCTGCTACT	TGG 1894
Qy	2101	AGGCATATCGACTACGCGATCATGGG	AGCAACACCCGCTCTGTGGATCTATGG	AATCTA 2160

1895	Db	AGCCACTATCGACTACGGGATCATGGCCACCACACCCCGTCTGTGGATCTATCGAATCTA	1954
2161	Qy	AATGTAAAGTTAAATCTCTAAATTAATTAATAAGTCCAGATTCTTCCATACGAACCTTAA	2220
1955	Db	AATGTAAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGAACCTTAA	2014
2221	Qy	CAGCAATTCCGGTAGCATCTAGACCTTCAACAGCAGCGCAGATCCATCACTCTCTGGCCAA	2280
2015	Db	CAGCAATTCCGGTAGCATCTAGACCTTCAACAGCAGCGCAGATCCATCACTCTCTGGCCAA	2074
2281	Qy	TATGTTTTCAGTCCCTTCAGGAGTTACGCTCTGTGAAGTGAATCAATCTTCGGAAGTTGCG	2340
2075	Db	TATGTTTTCAGTCCCTTCAGGAGTTACGCTCTGTGAAGTGAATCAATCTTCGGAAGTTGCG	2134
2341	Qy	TGTTAACTCCGCTGTATTGACGGGCATATCCGTACGTTTGGCAAAAGTGTGGTTGTTACCGG	2400
2135	Db	TGTTAACTCCGCTGTATTGACGGGCATATCCGTACGTTTGGCAAAAGTGTGGTTGTTACCGG	2194
2401	Qy	AGGAGTAAATCTCCAACAATCTCTCGAGAGTAGGCACCAACAAACACAGATCCACGGTGTT	2460
2195	Db	AGGAGTAAATCTCCAACAATCTCTCGAGAGTAGGCACCAACAAACACAGATCCACGGTGTT	2254
2461	Qy	GTACTTTGATCAACATAAGAAGAAGCATTTCTCGATTTGCAGGATCAAGTGTTCAGGAGCGT	2520
2255	Db	GTACTTTGATCAACATAAGAAGAAGCATTTCTCGATTTGCAGGATCAAGTGTTCAGGAGCGT	2314
2521	Qy	ACTGATTTGGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCGATAGGTTGTAGAGTGTG	2580
2315	Db	ACTGATTTGGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCGATAGGTTGTAGAGTGTG	2374
2581	Qy	CAATACACTTGGCTACAAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTGTGAAACAGCAT	2640
2375	Db	CAATACACTTGGCTACAAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTGTGAAACAGCAT	2434
2641	Qy	CTTCAATTTGGCAAGCTTCCTGTCTGTCTATCGACGCCAACAGAAATCACTCTGGGAAT	2700
2435	Db	CTTCAATTTGGCAAGCTTCCTGTCTGTCTATCGACGCCAACAGAAATCACTCTGGGAAT	2494
2701	Qy	CAATACCATGTTTCAGCTTGACACAGAAGGTTCTGAGGCAACGAATCTGGATCAGCGTATT	2760
2495	Db	CAATACCATGTTTCAGCTTGACACAGAAGGTTCTGAGGCAACGAATCTGGATCAGCGTATT	2554
2761	Qy	TATCAGCAATAACTAGAACTTTCAGAAGCCCGCAGCAGGCATCTCAATACTACACAGGGCTG	2820
2555	Db	TATCAGCAATAACTAGAACTTTCAGAAGCCCGCAGCAGGCATCTCAATACTACACAGGGCTG	2614
2821	Qy	ATGTGCTCAATTTGAACCATCATCTTTGGCAGCAGTAAACGAATCTGTTTCTCGAACCAATA	2880
2615	Db	ATGTGCTCAATTTGAACCATCATCTTTGGCAGCAGTAAACGAATCTGTTTCTCGAACCAATA	2674
2891	Qy	TTTTGTCACTTAGGAAACAGTTTCTGTTCCGTAAAGCCATAGCAGCTACTCCCTGGGCGC	2940
2675	Db	TTTTGTCACTTAGGAAACAGTTTCTGTTCCGTAAAGCCATAGCAGCTACTCCCTGGGCGC	2734
2941	Qy	CTCCTGTAGCACAGATACACTTTAGCACCACTTGTGGGCACCTAGATGACTTCTGGGG	3000
2735	Db	CTCCTGTAGCACAGATACACTTTAGCACCACTTGTGGGCACCTAGATGACTTCTGGGG	2794
3001	Qy	TAAAGGTTACCATCCTTCTTAGGTGGAGTGCAAAAACAAATTTCTTGCACACCAACCACTT	3060
2795	Db	TAAAGGTTACCATCCTTCTTAGGTGGAGTGCAAAAACAAATTTCTTGCACACCAACCACTT	2854
3061	Qy	TGGCAGGAACACCCAGCATCAGGGAAGTGGNAGCGAATTCGCGTTCCACCAGGAATAT	3120
2855	Db	TGGCAGGAACACCCAGCATCAGGGAAGTGGNAGCGAATTCGCGTTCCACCAGGAATAT	2914
3121	Qy	AGAGGCCAACTTTTCTCAATAGGTTCTGCMAAACGAGAGCAGACTACACAGGGCAAGTCT	3180
2915	Db	AGAGGCCAACTTTTCTCAATAGGTTCTGCMAAACGAGAGCAGACTACACAGGGCAAGTCT	2974
3181	Qy	CAACTTGCACAGCTCTCCGTTAGTTAGCTTCAATCGAATTTCTCGAGCTTATCTATAGAGA	3240

Db 2975 CAACCTTGCACAGTCTCGTTAGTTGAGCTTCATGGAATTCCTGACGTTATCTATAGAGA 3034  
 Qy 3241 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGCAATAGTTCCCTCTGGGAAAGGAGCTT 3300  
 Db 3035 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGCAATAGTTCCCTCTGGGAAAGGAGCTT 3094  
 Qy 3301 CTAACACAGGTGCTCTCAACGACCTCCATCAACCTTGGCAGTGTCTTAAAGGGCTT 3360  
 Db 3095 CTAACACAGGTGCTCTCAACGACCTCCATCAACCTTGGCAGTGTCTTAAAGGGCTT 3154  
 Qy 3361 TGTCAACATTTTGACGAACATTTGCGCAATTTGGTTGACTAATTCATAATCTGTTCGG 3420  
 Db 3155 TGTCAACATTTTGACGAACATTTGCGCAATTTGGTTGACTAATTCATAATCTGTTCGG 3214  
 Qy 3421 TTTTCTCGATAGACGACGAGGACATCTTCAATTTCTTGTGAGGAGGCTTAGAAGCTT 3480  
 Db 3215 TTTTCTCGATAGACGACGAGGACATCTTCAATTTCTTGTGAGGAGGCTTAGAAGCTT 3274  
 Qy 3481 CAATTTTGCACAAATCAATACGACCTTCAGAAGGACTTCTTTAGGTTTGGATTCTTCTT 3540  
 Db 3275 CAATTTTGCACAAATCAATACGACCTTCAGAAGGACTTCTTTAGGTTTGGATTCTTCTT 3334  
 Qy 3541 TAGGTTGTTCTGGTGATCTCGGCTTGGGATCTCTTCTTCTTCTAGTGACCTTTAGGG 3600  
 Db 3335 TAGGTTGTTCTGGTGATCTCGGCTTGGGATCTCTTCTTCTTCTAGTGACCTTTAGGG 3394  
 Qy 3601 ACTTCATATCCAGGTTCTCTCCACCTCTCCACGTCACACCGTACTTGGCACATCTAA 3660  
 Db 3395 ACTTCATATCCAGGTTCTCTCCACCTCTCCACGTCACACCGTACTTGGCACATCTAA 3454  
 Qy 3661 CTAATGCAAAATAAATAAGTACGACATCCCAAGGCTATATCTTCTTGGATTAGCTT 3720  
 Db 3455 CTAATGCAAAATAAATAAGTACGACATCCCAAGGCTATATCTTCTTGGATTAGCTT 3514  
 Qy 3721 CTGCAAGTTCAATCAGCTTCTCTCCATTTTAGGGTTCAACAAATCTCGTCTCAATA 3780  
 Db 3515 CTGCAAGTTCAATCAGCTTCTCTCCATTTTAGGGTTCAACAAATCTCGTCTCAATA 3574  
 Qy 3781 ACCGTTTGGTATAAGAACCTTCTGGAGCATTTCTTACGATCCCAAGGCTGGTTCCA 3840  
 Db 3575 ACCGTTTGGTATAAGAACCTTCTGGAGCATTTCTTACGATCCCAAGGCTGGTTCCA 3634  
 Qy 3841 TGGCTCTAAGACCTTTGATTGGCCAAAACAGGAGTGGCTTCAAGTGACAGAACCAA 3900  
 Db 3635 TGGCTCTAAGACCTTTGATTGGCCAAAACAGGAGTGGCTTCAAGTGACAGAACCAA 3694  
 Qy 3901 CACCTGTTTGTTCACCAACAAATTTCAAGCATCTCCATCACAATCCAATTCGATACCCA 3960  
 Db 3695 CACCTGTTTGTTCACCAACAAATTTCAAGCATCTCCATCACAATCCAATTCGATACCCA 3754  
 Qy 3961 GCACTTTTGTGTTGCTCCAGATAGACCTTTATACCAACCAACCTGACGAGATT 4020  
 Db 3755 GCACTTTTGTGTTGCTCCAGATAGACCTTTATACCAACCAACCTGACGAGATT 3814  
 Qy 4021 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4080  
 Db 3815 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 3874  
 Qy 4081 GGGCCACAGTAATFAGAAGAGTACGCCACAAAGTAGTGAATAGACCATCGGGGGCTT 4140  
 Db 3875 GGGCCACAGTAATFAGAAGAGTACGCCACAAAGTAGTGAATAGACCATCGGGGGCTT 3934  
 Qy 4141 CAGTAGTCAAGACGCCAACAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAA 4200  
 Db 3935 CAGTAGTCAAGACGCCAACAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAA 3994  
 Qy 4201 GTTCTGATTTAGTAGTCAATTTGCCGAGCATCAATAATGGGATTATACAGGAACACAG 4260  
 Db 3995 GTTCTGATTTAGTAGTCAATTTGCCGAGCATCAATAATGGGATTATACAGGAACACAG 4054  
 Qy 4261 TGGAGTCACTACCAACTTTGGGTTCTCAGAAAAGCATATAACAGTTCTTACTACCGC 4320  
 Db 4055 TGGAGTCACTACCAACTTTGGGTTCTCAGAAAAGCATATAACAGTTCTTACTACCGC 4114

Qy 4321 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCATACTAGCAT 4380  
 Db 4115 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCATACTAGCAT 4174  
 Qy 4381 TAGCGGCAAGGATGCAACTTTTATCAACCGAGGTCCTATAGATAAACCCTAGCCCTGGGA 4440  
 Db 4175 TAGCGGCAAGGATGCAACTTTTATCAACCGAGGTCCTATAGATAAACCCTAGCCCTGGGA 4234  
 Qy 4441 TCATCTTTTGACAACTCTTTCTGCCAAATCTAGGTCGAAATCACCTTCAATTTGATACCAT 4500  
 Db 4235 TCATCTTTTGACAACTCTTTCTGCCAAATCTAGGTCGAAATCACCTTCAATTTGATACCAT 4294  
 Qy 4501 TATTGTCAACTTGGACAAAGTTGTCGATCAGCTCCTCAAAATTTGGTCTCTGTAAACGGATG 4560  
 Db 4295 TATTGTCAACTTGGACAAAGTTGTCGATCAGCTCCTCAAAATTTGGTCTCTGTAAACGGATG 4354  
 Qy 4561 ACTCAACTTGCACATTTAACTTGAAGCTCAGTCGATTTGAGTGAATTCGATCAGGTTGTGCA 4620  
 Db 4355 ACTCAACTTGCACATTTAACTTGAAGCTCAGTCGATTTGAGTGAATTCGATCAGGTTGTGCA 4414  
 Qy 4621 GCTGTCAGCAGCATAGGGAACACGGCTTTTCTACCAAACTCAAGGAATTTATCAAACT 4680  
 Db 4415 GCTGTCAGCAGCATAGGGAACACGGCTTTTCTACCAAACTCAAGGAATTTATCAAACT 4474  
 Qy 4681 CTGCAACTTGGCGTATGCGAGTAGCAAGGAATGTCATCTTGAAGTCGACAGTGAG 4740  
 Db 4475 CTGCAACTTGGCGTATGCGAGTAGCAAGGAATGTCATCTTGAAGTCGACAGTGAG 4534  
 Qy 4741 TGTAGCTTGGAGAAATCTGAAGCGCTATTTTATTTATCAGTGAGTCAGTCATCAGGAGA 4800  
 Db 4535 TGTAGCTTGGAGAAATCTGAAGCGCTATTTTATTTATCAGTGAGTCAGTCATCAGGAGA 4594  
 Qy 4801 TCCTCTACCGCGGACGATCTGCGGA----- 4828  
 Db 4595 TCCTCTACCGCGGACGATCTGCGCGGACCTCGAGGGGGGGGGGGGGGGCTGCTG 4654  
 Qy 4829 ----- 4828  
 Db 4655 CTCGTGAAGAGGTGTTGCTGACTCATACAGGCTTGATCGCCCATCATCCAGCCAGA 4714  
 Qy 4829 ----- 4828  
 Db 4715 AAGTGAGGAGCCAGGTTGATGAGAGCTTTGTTGTTAGGTGGACAGTGGTGTATTGA 4774  
 Qy 4829 ----- 4828  
 Db 4775 ACTTTTGTGTTGCCACGGAACGGTCTGCGTGTGGGAAGATGCGTGATCTGATCCTTCA 4834  
 Qy 4829 ----- 4828  
 Db 4835 ACTCAGCAAAAGTTGCAATTTATTCACAAAGCGCGCTCCCGTCAAGTCAGCGTAATGCT 4894  
 Qy 4829 ----- 4828  
 Db 4895 CTGCCAGTGTTAACCAATTAACCAATTTCTGATTAGAAAAAACTCATCGAGCATCAAAATG 4954  
 Qy 4829 ----- 4828  
 Db 4955 AAACCTGCATTTTATTCATATCAGGATTATCAATACCATATTTTGAAGAAAGCGGTTTCTG 5014  
 Qy 4829 ----- 4828  
 Db 5015 TAATGAAGAGGAAAAACTCACCGAGGAGTTCATAGTAGGCAAGATCCTGGTATCGGTC 5074  
 Qy 4829 ----- 4828  
 Db 5075 TGGGATTCGACTCGTCCACATCAATACACCTATTATTTCCCTCGTCAAAAAAATAAG 5134  
 Qy 4829 ----- 4828  
 Db 5135 GTTATCAAGTGAGAAATCACCATGAGTCAGCGACTGAATCCGGTGAGATGGCAAAAGCTT 5194

QY 4829 ----- 4828  
Db 5195 ATGCATTTTCCAGACTTGTTCACAGGCGAGCCATTAGCTCGTCATCAAAATCACT 5254  
QY 4829 ----- 4828  
Db 5255 CGCATCAACCAACCGTTATTTCATTCGTGATTTGCCCTGAGGAGACGAATAACGGATC 5314  
QY 4829 ----- 4828  
Db 5315 GCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGAACACTGCCAG 5374  
QY 4829 ----- 4828  
Db 5375 GCATCAACAATATTTTCACTGAATCAGGATATTTCTTAATACCTGGAAATGCTGTTT 5434  
QY 4829 ----- 4828  
Db 5435 CCCGGGATCGAGTGTGAGTAACCAATGCATCATCAGGACTACGGATAAATGCTTGAT 5494  
QY 4829 ----- 4828  
Db 5495 GGTGGAAAGGCATAAAATCCGTGAGCCAGTTTGTCTGACCATCTCATCTGTAACATC 5554  
QY 4829 ----- 4828  
Db 5555 ATTGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGATCGGCTTCCCATA 5614  
QY 4829 ----- 4828  
Db 5615 CAATCGATAGATTGTCGACCTGATTTGCCGACATTTATCGGAGCCCATTTATACCCATA 5674  
QY 4829 ----- 4828  
Db 5675 TAAATCAGATCCATGTTGGAAATTTAATCGCGGCTCGAGCAAGCTTCCCGTTGAAT 5734  
QY 4829 ----- 4828  
Db 5735 ATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTATTGTTTCATGA 5794  
QY 4829 ----- 4828  
Db 5795 TGATATATTTTATCTTGTGCAATGTAAATCAGAGATTTTGTAGACACAAGTGGCTTTC 5854  
QY 4829 ----- 4879  
Db 5855 CCCCCCCCCCTGCAGGTCGACATCACCGGCGCCACAGGTGCGGTTGCTGGGCGCTATAT 5914  
QY 4880 CGCGACATACCGATGGGGAAGATCGGCTCGCCACTTTGGGCTCATGAGCGCTTGT 4939  
Db 5915 CGCGACATACCGATGGGGAAGATCGGCTCGCCACTTTGGGCTCATGAGCGCTTGT 5974  
QY 4940 CGGCTGGGTATGTTGGCAGCCCGGCGGCGGAGCTGTTGGCGCCATCTCTTTCGA 4999  
Db 5975 CGCGTGGGTATGTTGGCAGCCCGGCGGCGGAGCTGTTGGCGCCATCTCTTTCGA 6034  
QY 5000 TGACCAATTCCTTGGCGGCGGCTGCTCAACGGGCTCAACCTACTACTGGGCTGCTTCCT 5059  
Db 6035 TGACCAATTCCTTGGCGGCGGCTGCTCAACGGGCTCAACCTACTACTGGGCTGCTTCCT 6094  
QY 5060 ANTGCAGAGTCGATAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAAATGGT 5119  
Db 6095 AATGCAGAGTCGATAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAAATGGT 6154  
QY 5120 GATACCCGCATTCCTCAGTGTCTTGAGTCTTCCTATCAGATTATGCCCAACTAAAGCAAC 5179  
Db 6155 GATACCCGCATTCCTCAGTGTCTTGAGTCTTCCTATCAGATTATGCCCAACTAAAGCAAC 6214  
QY 5180 CGGAGGAGGATTTCAATGTAATTTCTCTGACTTTTGTGTCATGACTCGAACTG 5239  
Db 6215 CGGAGGAGGATTTCAATGTAATTTCTCTGACTTTTGTGTCATGACTCGAACTG 6274  
QY 5240 TGAGACTATCTCGGTATGACAGCAGAAATGTCCTTCTTGGAGACAGTAATGAGTCCC 5299

Db 6275 TGAGACTATCTCGGTATGACAGCAAAATGTCCTTCTTGGACAGTAATGAAGTCCC 6334  
QY 5300 ACCAATAAAGAAATCCTTGTATCAGGAACAACCTTCTTGTTCGAACCTTTTCGGTGCC 5359  
Db 6335 ACCAATAAAGAAATCCTTGTATCAGGAACAACCTTCTTGTTCGAACCTTTTCGGTGCC 6394  
QY 5360 TTGAACCTATAAAATGTAGAGTGGATATGTCGGGTAGAAATGGAGCGGCAAAATGCTTACC 5419  
Db 6395 TTGAACCTATAAAATGTAGAGTGGATATGTCGGGTAGAAATGGAGCGGCAAAATGCTTACC 6454  
QY 5420 TTCTGACCTTCAAGAGGTATGTAGGGTTGTAGATACCTGATGTCGCAACTTCAGTGACAAC 5479  
Db 6455 TTCTGACCTTCAAGAGGTATGTAGGGTTGTAGATACCTGATGTCGCAACTTCAGTGACAAC 6514  
QY 5480 GTTGCTATTTCGTTCAAAACCAATCCGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTAT 5539  
Db 6515 GTTGCTATTTCGTTCAAAACCAATCCGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTAT 5574  
QY 5540 GATCCAAAGCAGTCGGCTCTTGAACCTGACATAGTGTCTGCTGTTTGTAGGTCACTTT 5599  
Db 6575 GATCCAAAGCAGTCGGCTCTTGAACCTGACATAGTGTCTGCTGTTTGTAGGTCACTTT 6634  
QY 5600 TGTATCAATAAATCTAGTCTTTGATCTAAATCTTGAAGCAAGCCAAAGCGATAAATACC 5659  
Db 6635 TGTATCAATAAATCTAGTCTTTGATCTAAATCTTGAAGCAAGCCAAAGCGATAAATACC 5694  
QY 5660 CAAATCTAAACCTTTTAAACGTTTAAAGCAAGTATGTCGCTGTATTAACCCCC 5719  
Db 6695 CAAATCTAAACCTTTTAAACGTTTAAAGCAAGTATGTCGCTGTATTAACCCCC 6754  
QY 5720 AATCAGCTCGTAGTCTGATCTCACTCACTTGAAGGCGCACTATCTTGTGTTTAGAGAAAT 5779  
Db 6755 AATCAGCTCGTAGTCTGATCTCACTCACTTGAAGGCGCACTATCTTGTGTTTAGAGAAAT 6814  
QY 5780 TTGCGGAGATCGGATATCGAGAAAGGATGACCTGATTTTAAACGTTGAATTTATCTCAA 5839  
Db 6815 TTGCGGAGATCGGATATCGAGAAAGGATGACCTGATTTTAAACGTTGAATTTATCTCAA 6874  
QY 5840 GATCTCTGCTCGCGGTTTTCGTTGATGACGTTGAAACCTCTGACATGACGCTCCCG 5899  
Db 6875 GATCTCTGCTCGCGGTTTTCGTTGATGACGTTGAAACCTCTGACATGACGCTCCCG 6934  
QY 5900 GAGACGCTCACAGCTTGTCTGTAAGCGGATCGCGGAGCAGAAAGCCGTCAGGCGCG 5959  
Db 5935 GAGACGCTCACAGCTTGTCTGTAAGCGGATCGCGGAGCAGAAAGCCGTCAGGCGCG 5994  
QY 5960 TCAGCGGCTTGGCGGTTGTCGGGCGCAGCCATGACCCAGTCACTAGCAGTACGGA 6019  
Db 6995 TCAGCGGCTTGGCGGTTGTCGGGCGCAGCCATGACCCAGTCACTAGCAGTACGGA 7054  
QY 6020 GTGTATCTGCTTAACTATGCGGCATCAGACAGATTTGTACTGAGGTGACCATATGC 6079  
Db 7055 GTGTATCTGCTTAACTATGCGGCATCAGACAGATTTGTACTGAGGTGACCATATGC 7114  
QY 6080 GTGTGAAATACCGCACAGATCGTAAGGAGAAATACCGCATCAGGCGCTCTTCCGCTT 6139  
Db 7115 GTGTGAAATACCGCACAGATCGTAAGGAGAAATACCGCATCAGGCGCTCTTCCGCTT 7174  
QY 6140 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 6199  
Db 7175 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 7234  
QY 6200 CAAAGCGGTAATACGTTTATCCAGAAATCAGGAGTAAACCGCAGAAAGAAATGTAG 6259  
Db 7235 CAAAGCGGTAATACGTTTATCCAGAAATCAGGAGTAAACCGCAGAAAGAAATGTAG 7294  
QY 6260 CAAAGCGGTAATACGTTTATCCAGAAATCAGGAGTAAACCGCAGAAAGAAATGTAG 6319  
Db 7295 CAAAGCGGTAATACGTTTATCCAGAAATCAGGAGTAAACCGCAGAAAGAAATGTAG 7354  
QY 6320 GCTCGGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 6379



7355 GGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 7414  
 6380 CGACAGACTATAAAGATACAGCGCTTCCCTCGAAGCTCCCTCGCGCTCTCCCTG 6439  
 7415 CGACAGACTATAAAGATACAGCGCTTCCCTCGAAGCTCCCTCGCGCTCTCCCTG 7474  
 6440 TTCCGACCTCCGCGCTTACCGGATACCTGTCGCTTTTCCCTTCGGGAAGCGTGGCGC 6499  
 7475 TTCCGACCTCCGCGCTTACCGGATACCTGTCGCTTTTCCCTTCGGGAAGCGTGGCGC 7534  
 6500 TTCTCTAGCTACGCTGTAGGTATCTCAGTTCGGTTCGGTTCGCTCCAGCTGG 6559  
 7535 TTCTCTAGCTACGCTGTAGGTATCTCAGTTCGGTTCGGTTCGCTCCAGCTGG 7594  
 6560 GCTGTGTCAGAACCCCGCTTCAGCCGACCGCTTCGCGCTTATCCGGTAACTATCGTC 6619  
 7595 GCTGTGTCAGAACCCCGCTTCAGCCGACCGCTTCGCGCTTATCCGGTAACTATCGTC 7654  
 6620 TTGAGTCCAAACCCCGGTAAAGACAAGCTTATCGGACCTGGCAGCAGCACTGGTAAACAGGA 6679  
 7655 TTGAGTCCAAACCCCGGTAAAGACAAGCTTATCGGACCTGGCAGCAGCACTGGTAAACAGGA 7714  
 6680 TTACGACGCGAGGTATGTAGCGGTCTACAGATTCCTTGAAGTGGTGGCCCTAACTACG 6739  
 7715 TTACGACGCGAGGTATGTAGCGGTCTACAGATTCCTTGAAGTGGTGGCCCTAACTACG 7774  
 6740 GCTACACTAGAACGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTATACCTCGGAA 6799  
 7775 GCTACACTAGAACGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTATACCTCGGAA 7834  
 6800 AAAGAGTGGTAGCTCTGATCCGCGAACAACACCGCTGTAGCGGTGGTTTTTTT 6859  
 7835 AAAGAGTGGTAGCTCTGATCCGCGAACAACACCGCTGTAGCGGTGGTTTTTTT 7894  
 6860 TTTCGAAGCAGCAGATTCAGCGCAGAAAAAAGGATCTCAAGAGATTCCTTTGATCTTTT 6919  
 7895 TTTCGAAGCAGCAGATTCAGCGCAGAAAAAAGGATCTCAAGAGATTCCTTTGATCTTTT 7954  
 6920 CTACGGGTCTGAGCTCAGTGGAGCAAGAACTCAGCTTAAGGATTTTGGTCTATGAGAT 6979  
 7955 CTACGGGTCTGAGCTCAGTGGAGCAAGAACTCAGCTTAAGGATTTTGGTCTATGAGAT 8014  
 6980 TATCAAAAAGGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCT 7039  
 8015 TATCAAAAAGGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCT 8074  
 7040 AAAGTATATAGTAAACTTTGGTCTGACAGTACCAATGTTAATCAGTGGACCTTA 7099  
 8075 AAAGTATATAGTAAACTTTGGTCTGACAGTACCAATGTTAATCAGTGGACCTTA 8134  
 7100 TCTCAGCGATCTGCTATTTTCTGTTTATCCATCTGACTCCCGCTGCTGATAGATAA 7159  
 8135 TCTCAGCGATCTGCTATTTTCTGTTTATCCATCTGACTCCCGCTGCTGATAGATAA 8194  
 7160 CTACGATACGGAGGGCTTACCATCTGCGCCGATGCTGCAATGATACCGGAGACCCAC 7219  
 8195 CTACGATACGGAGGGCTTACCATCTGCGCCGATGCTGCAATGATACCGGAGACCCAC 8254  
 7220 GCTACCGGCTCCAGATTTATCAGCAATAAACAGCAGCGGAGGCGGAGCGAGAA 7279  
 8255 GCTACCGGCTCCAGATTTATCAGCAATAAACAGCAGCGGAGGCGGAGCGAGAA 8314  
 7280 GTGGTCTCGCACTTTATCCGCTCCATCCAGTCTTAAATTTGTTCGGGAAGCTAGAG 7339  
 8315 GTGGTCTCGCACTTTATCCGCTCCATCCAGTCTTAAATTTGTTCGGGAAGCTAGAG 8374  
 7340 TAAGTAGTTCGGCAGTAAATAGTTTGGCAAGTGTGGCCATGCTGCGAGCATCTGG 7399  
 8375 TAAGTAGTTCGGCAGTAAATAGTTTGGCAAGTGTGGCCATGCTGCGAGCATCTGG 8434  
 7400 TGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCGCAACGATCAAGGGCAG 7459  
 8435 TGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCGCAACGATCAAGGGCAG 8494

7460 TTACATGATCCCCCATGTTTGTGCAAAAAAGCGGTAGTCTCTTCGGTCCCTCCGATGTTG 7519  
 8495 TTACATGATCCCCCATGTTTGTGCAAAAAAGCGGTAGTCTCTTCGGTCCCTCCGATGTTG 8554  
 7520 TCAGAAGTAAAGTTTGGCGCAGTGTATCACTCATGTTATGCGAGCAGCTGCAATATCTC 7579  
 8555 TCAGAAGTAAAGTTTGGCGCAGTGTATCACTCATGTTATGCGAGCAGCTGCAATATCTC 8614  
 7580 TTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAT 7639  
 8615 TTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAT 8674  
 7640 TCTGAGTATAGTGTATGGCGGACCGAGTGTCTCTTCGCCGCGCTCAACCGGATATA 7699  
 8675 TCTGAGTATAGTGTATGGCGGACCGAGTGTCTCTTCGCCGCGCTCAACCGGATATA 8734  
 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATATTGGAACAGTTCTTCGGGGCGAA 7759  
 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATATTGGAACAGTTCTTCGGGGCGAA 8794  
 7760 AACTCTCAGGATCTTACCGCTGTGAGTCCAGTTCGATGTAAACCCACTGTCACCCCA 7819  
 8795 AACTCTCAGGATCTTACCGCTGTGAGTCCAGTTCGATGTAAACCCACTGTCACCCCA 8854  
 7820 ACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTCTGGGTGAGCAAAAACAGGAAGGC 7879  
 8855 ACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTCTGGGTGAGCAAAAACAGGAAGGC 8914  
 7880 AAATGCGCGAAAAAAGGGAATTAAGGCGGACACCGAATGTTGAATACTCATCTCTCC 7939  
 8915 AAATGCGCGAAAAAAGGGAATTAAGGCGGACACCGAATGTTGAATACTCATCTCTCC 8974  
 7940 TTTTTCATATATTATTGAAGCATTTTATCAGGCTTATTTCTCATGCGGATACATATTTG 7999  
 8975 TTTTTCATATATTATTGAAGCATTTTATCAGGCTTATTTCTCATGCGGATACATATTTG 9034  
 8000 AATGATTTAGAAAAATAAACAATAGGGGTTCGGCGACATTTCCCGAAAAAGTGCAC 8059  
 9035 AATGATTTAGAAAAATAAACAATAGGGGTTCGGCGACATTTCCCGAAAAAGTGCAC 9094  
 8060 CTGAGCTCTAGAACCAATATTATCATGACATTAACCTATAAATAAGGCTATCAGA 8119  
 9095 CTGAGCTCTAGAACCAATATTATCATGACATTAACCTATAAATAAGGCTATCAGA 9154  
 8120 GGCCCTTTCTGCTTCAAGAAATAATTCTCATGTTTGAAGCTTATCATCGATAAGTGCAC 8179  
 9155 GGCCCTTTCTGCTTCAAGAAATAATTCTCATGTTTGAAGCTTATCATCGATAAGTGCAC 9214  
 8180 TCATGTTGTTTGAATAGACGAGTCCGGAACACTGAAAAATAACAGTTTATTT 8239  
 9215 TCATGTTGTTTGAATAGACGAGTCCGGAACACTGAAAAATAACAGTTTATTT 9274  
 8240 CG 8241  
 9275 CG 9276

RESULT 3  
 US-10-302-840A-13  
 ; Sequence 13, Application US/10302840A  
 ; Publication No. US20030134794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Madison, Edwin L.  
 ; APPLICANT: Ong, Edgar O.  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING SERINE PROTEASE CVSP17, THE ENCODED  
 ; TITLE OF INVENTION: POLYPEPTIDES AND METHODS BASED THEREON  
 ; FILE REFERENCE: 24745-1622  
 ; CURRENT APPLICATION NUMBER: US/10/302,840A  
 ; PRIOR FILING DATE: 2003-01-24  
 ; PRIOR APPLICATION NUMBER: 60/332,015  
 ; PRIOR FILING DATE: 2001-11-20  
 ; NUMBER OF SEQ ID NOS: 18



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
US-10-302-840A-13

Query Match      79.6%; Score 6561.6; DB 15; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY 1 AGATCTAACATCCAAAGAGCAAGAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
DB 1 AGATCTAACATCCAAAGAGCAAGAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
QY 61 GTCCATTCTCACATAGTSCCAACCGCAACAGAGGGGATACACTAGCAGACCGT 120
DB 61 GTCCATTCTCACATAGTSCCAACCGCAACAGAGGGGATACACTAGCAGACCGT 120
QY 121 TGCAGAGCAGAGCTCCACTCCTCTCTCTCAACACCCACTTTTCCCATCGAAACCC 180
DB 121 TGCAGAGCAGAGCTCCACTCCTCTCTCTCAACACCCACTTTTCCCATCGAAACCC 180
QY 181 AGCCAGGTTATTGGGCTTGATTGGAGCTCGCTCATTTCCAAATTCCTTATTAGGCTACTA 240
DB 181 AGCCAGGTTATTGGGCTTGATTGGAGCTCGCTCATTTCCAAATTCCTTATTAGGCTACTA 240
QY 241 ACACATGACTTTATTAGCTGTCTATCTGCGCCCGCTGGCAGAGTTCAITGTTGTTA 300
DB 241 ACACATGACTTTATTAGCTGTCTATCTGCGCCCGCTGGCAGAGTTCAITGTTGTTA 300
QY 301 TTTCCGAATCAACAGCTCGCATTTACACCCGAACATCACTCCAGATGAGGGCTTTCTG 360
DB 301 TTTCCGAATCAACAGCTCGCATTTACACCCGAACATCACTCCAGATGAGGGCTTTCTG 360
QY 361 AGTGTGGGTCAATAGTTTCATGTTCCCAAAATGGGCCAAAACCTGACAGTTTAAACGCT 420
DB 361 AGTGTGGGTCAATAGTTTCATGTTCCCAAAATGGGCCAAAACCTGACAGTTTAAACGCT 420
QY 421 GTCTTGGAACTATATGACAAAGAGCGTGATCTCATCCAGATCAATTAAGTTGGTTCG 480
DB 421 GTCTTGGAACTATATGACAAAGAGCGTGATCTCATCCAGATCAATTAAGTTGGTTCG 480
QY 481 TTGAAATGCTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCCATACCGTTGT 540
DB 481 TTGAAATGCTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCCATACCGTTGT 540
QY 541 CTGTTTGGTATTGATTGACCAATGCTCAAAATTAATCTCATTAATGCTTAGCGCAGTCT 600
DB 541 CTGTTTGGTATTGATTGACCAATGCTCAAAATTAATCTCATTAATGCTTAGCGCAGTCT 600
QY 601 CTCTATCGCTTCTGNAACCCCGGTGCACCTGTGCCGAAACGCAATGGGAAACACCCGCT 660
DB 601 CTCTATCGCTTCTGNAACCCCGGTGCACCTGTGCCGAAACGCAATGGGAAACACCCGCT 660
QY 661 TTTTGGATGATATGCAATGTCTCCACATTTGATGCTTCCAAAGATTTCTGTGGGAACTACT 720
DB 661 TTTTGGATGATATGCAATGTCTCCACATTTGATGCTTCCAAAGATTTCTGTGGGAACTACT 720
QY 721 GCTGATGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGCAGCAAT 780
DB 721 GCTGATGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGCAGCAAT 780
QY 781 ATATAACAGAGGAGCTGCCCTGTCTTAAACCTTTTTTTTATCATCATTAATAGCTT 840
DB 781 ATATAACAGAGGAGCTGCCCTGTCTTAAACCTTTTTTTTATCATCATTAATAGCTT 840
QY 841 ACTTTCATTAATGCGACTGGTTCGAATGCAAGCTTTTGATTTTAAGGACTTTTAACGA 900
DB 841 ACTTTCATTAATGCGACTGGTTCGAATGCAAGCTTTTGATTTTAAGGACTTTTAACGA 900
QY 901 CAACCTTGAGAGATCAAAAACCACTAATTAATTCGAGGATCCAAAGATGAGATTTTCT 960
DB 901 CAACCTTGAGAGATCAAAAACCACTAATTAATTCGAGGATCCAAAGATGAGATTTTCT 960

901 CAACCTTGAGAGATCAAAAACCACTAATTAATTCGAGGATCCAAAGATGAGATTTCT 960
961 TCAATTTTACTGCAAGTTTATTCGACATCTCTCCGATAGCTGCTCCAGTCAACT 1020
961 TCAATTTTACTGCAAGTTTATTCGACATCTCTCCGATAGCTGCTCCAGTCAACT 1020
1021 ACACAGAGATGAACCGCACAATTCGGCTGAGCTGTCATCGTGTACTCAGATTTA 1080
1021 ACACAGAGATGAACCGCACAATTCGGCTGAGCTGTCATCGTGTACTCAGATTTA 1080
1081 GAAGGGATTTTCGATTTGCTGTTTGGCATTTTCCAAACAGCACAATAACGGTTATTG 1140
1081 GAAGGGATTTTCGATTTGCTGTTTGGCATTTTCCAAACAGCACAATAACGGTTATTG 1140
1141 TTTTAAATACACTATTGCGCAGCATTTGCTCTAAAGAGAGGGTATCTCTCAGAAA 1200
1141 TTTTAAATACACTATTGCGCAGCATTTGCTCTAAAGAGAGGGTATCTCTCAGAAA 1200
1201 AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGAAAGAAACATGTATCATG 1260
1201 AGAGAGGCTGAAGCTTA----- 1217
1261 GCGCCCGGAGAGGANTAAATTTGGCTTCCCGGCTGTCCCGCCAGCAGTGCAGGAG 1320
1218 ----- 1217
1321 AGAGGTTGCTGTTTGTATGACAGTGTCCGGGGATTTCCCGTGTCTTCCACCCCATGGCC 1380
1218 ----- 1217
1381 ATCGAACAACACTCAAGAAGAAGATGTCCCTTCTAACTAGTGGCGTAGAATTTCCCTAGG 1440
1218 -----CGTAGAATTTCCCTAGG 1234
1441 CGCCCGGAAATTAATTCGCTTTAGACATGACTGTTTCCCTCAGTTCAAGTTGGGCACCTAGC 1500
1235 CGCCCGGAAATTAATTCGCTTTAGACATGACTGTTTCCCTCAGTTCAAGTTGGGCACCTAGC 1294
1501 AGAAGACCCGCTCTGTAGATTTCTAATCAAGAGGATGTCAAGATGCCATTTCCCTGAGAG 1560
1295 AGBAGACCCGCTCTGTAGATTTCTAATCAAGAGGATGTCAAGATGCCATTTCCCTGAGAG 1354
1561 ATGACGCTTCATTTTGTATCTTTTATTTTATTTGTAACCTATATAGTATAGGATTTTTTT 1620
1355 ATGACGCTTCATTTTGTATCTTTTATTTTATTTGTAACCTATATAGTATAGGATTTTTTT 1414
1621 GTCATTTTCTTCTCTGACAGCTTCTCTGATCAGCCTATCTCCAGCTGATGAA 1680
1415 GTCATTTTCTTCTCTGACAGCTTCTCTGATCAGCCTATCTCCAGCTGATGAA 1474
1681 TATCTTTGCTAGGGTTTGGGAAATCATTCGAGTTTGTATGTTTTTTCTTGGTATTTTCCC 1740
1475 TATCTTTGCTAGGGTTTGGGAAATCATTCGAGTTTGTATGTTTTTCTTGGTATTTCCC 1534
1741 ACTCCTCTTCAGAGTACAGAAGATTAAAGTGAAGTTGCTTTGTCAAGCTTATCGATAA 1800
1535 ACTCCTCTTCAGAGTACAGAAGATTAAAGTGAAGTTGCTTTGTCAAGCTTATCGATAA 1594
1801 GCTTTAATCGGTAGTTTATCAGTTAAATTTGCTAAACGAGTCAAGGACCCGCTGATGAA 1860
1595 GCTTTAATCGGTAGTTTATCAGTTAAATTTGCTAAACGAGTCAAGGACCCGCTGATGAA 1654
1861 AUCTAAACATGCGCTCATCTCGGACCGCTCACCCCTGGATGCTGTAGGCAATAG 1920
1655 AUCTAAACATGCGCTCATCTCGGACCGCTCACCCCTGGATGCTGTAGGCAATAG 1714
1921 CTTGGTTATGCGGTACTCTCGGCGCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1980
1715 CTTGGTTATGCGGTACTCTCGGCGCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1774
1981 CAGTCACTATGCGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCT 2040
1775 CAGTCACTATGCGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCT 1834
```

QY 2041 TCTCGGAGCACTGTCCGACCGCTTTGGCCGCCGCCAGTCTCTGCTCGCTCGCTACTTGG 2100  
Db 1835 TCTCGGAGCACTGTCCGACCGCTTTGGCCGCCGCCAGTCTCTGCTCGCTCGCTACTTGG 1894  
QY 2101 AGCCACTATCGACTACCGCATCATGGCGACCAACACCCGCTCTCTGGATCTATCGAATCTA 2160  
Db 1895 AGCCACTATCGACTACCGCATCATGGCGACCAACACCCGCTCTCTGGATCTATCGAATCTA 1954  
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220  
Db 1955 AATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2014  
QY 2221 CAGCAATTCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2280  
Db 2015 CAGCAATTCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2074  
QY 2281 TATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTGATGAACCTTCTGGAAGGTTGCAG 2340  
Db 2075 TATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTGATGAACCTTCTGGAAGGTTGCAG 2134  
QY 2341 TGTTAACTCCGCTGTATGAGCGGCATATCGGTAGCTTGGCAAGTGTGTTGGTACCCG 2400  
Db 2135 TGTTAACTCCGCTGTATGAGCGGCATATCGGTAGCTTGGCAAGTGTGTTGGTACCCG 2194  
QY 2401 AGGAGTAATCTCCAACTCTCTGGAGTAGGCACCAACAAACACAGATCCAGCGTGT 2460  
Db 2195 AGGAGTAATCTCCAACTCTCTGGAGTAGGCACCAACAAACACAGATCCAGCGTGT 2254  
QY 2461 GTACTGATCAACATAAGAGAGCATCTCGATTTGCAGGATCAAGTGTTCAGGAGCGT 2520  
Db 2255 GTACTGATCAACATAAGAGAGCATCTCGATTTGCAGGATCAAGTGTTCAGGAGCGT 2314  
QY 2521 ACTGATTTGGCAATTTCCAAAGCCCTGCTAGGTTGCAACCGATAGGTTGTAGAGTG 2580  
Db 2315 ACTGATTTGGCAATTTCCAAAGCCCTGCTAGGTTGCAACCGATAGGTTGTAGAGTG 2374  
QY 2581 CAATACTTGGTACCAATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAACAGCAT 2640  
Db 2375 CAATACTTGGTACCAATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAACAGCAT 2434  
QY 2641 CTTCATTTCTGGAGCTCTCTGCTGCTATATCGACAGCAACAGATCACTCGGCAAT 2700  
Db 2435 CTTCATTTCTGGAGCTCTCTGCTGCTATATCGACAGCAACAGATCACTCGGCAAT 2494  
QY 2701 CAATACCACTTTCAGCTTGAGACAGAGGCTTGAGGCAACCAATCTGGATCAGCGTATT 2760  
Db 2495 CAATACCACTTTCAGCTTGAGACAGAGGCTTGAGGCAACCAATCTGGATCAGCGTATT 2554  
QY 2761 TATCAGCAATTAACATAGCACTTCAAGGCGCCAGCGCAGTGTCAATTAACAGGCGCTG 2820  
Db 2555 TATCAGCAATTAACATAGCACTTCAAGGCGCCAGCGCAGTGTCAATTAACAGGCGCTG 2614  
QY 2821 ATGTGTCAATTTGAACCATCATCTTGGCAGCAGTAACGAACCTGTTTCTTGGACCAATA 2880  
Db 2615 ATGTGTCAATTTGAACCATCATCTTGGCAGCAGTAACGAACCTGTTTCTTGGACCAATA 2674  
QY 2881 TTTTGTCACTTAGGAACAGTTCCTGTTCCGTTAAGCCATAGAGCTACTGCTGGCGC 2940  
Db 2675 TTTTGTCACTTAGGAACAGTTCCTGTTCCGTTAAGCCATAGAGCTACTGCTGGCGC 2734  
QY 2941 CTCCTGCTAGCAGATACACTTAGCACAACCTTGTGGCAACCTAGATGACTTCTGGGG 3000  
Db 2735 CTCCTGCTAGCAGATACACTTAGCACAACCTTGTGGCAACCTAGATGACTTCTGGGG 2794  
QY 3001 TAAGGTTACCATCTCTTCTAGTGGAGATGCAAAAACAAATTTCTTGGCAACAGCACTT 3060  
Db 2795 TAAGGTTACCATCTCTTCTAGTGGAGATGCAAAAACAAATTTCTTGGCAACAGCACTT 2854  
QY 3061 TGGCAGGAACACCCAGCATCAGGGAAGTGGAGGAGCAATTTGGGTTCCACAGGAAATAT 3120  
Db 2855 TGGCAGGAACACCCAGCATCAGGGAAGTGGAGGAGCAATTTGGGTTCCACAGGAAATAT 2914

QY 3121 AGAGGCCAACTTTCTCAATAGTCTTGTCAAAACGAGAGCAGACTTACACAGGCGCAAGTCT 3180  
Db 2915 AGAGGCCAACTTTCTCAATAGTCTTGTCAAAACGAGAGCAGACTTACACAGGCGCAAGTCT 2974  
QY 3181 CAACCTTGCAACGCTCTCGTTAGTTAGCTTCATGGAAATTTCTTGACGCTTATCTATAGAGA 3240  
Db 2975 CAACCTTGCAACGCTCTCGTTAGTTAGCTTCATGGAAATTTCTTGACGCTTATCTATAGAGA 3034  
QY 3241 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCATAAGTTCCTCTGGGAAAGGAGCTT 3300  
Db 3035 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCATAAGTTCCTCTGGGAAAGGAGCTT 3094  
QY 3301 CTAAACAGAGTGTCTTCAAGCGACTCATCAAACTTTGGGAGGCTTATAGAAAGCT 3360  
Db 3095 CTAAACAGAGTGTCTTCAAGCGACTCATCAAACTTTGGGAGGCTTATAGAAAGCT 3154  
QY 3361 TGTCAACATTTTGAGCAACATTTGCGAATTTGACTTAATTCATAATCTGTTCCG 3420  
Db 3155 TGTCAACATTTTGAGCAACATTTGCGAATTTGACTTAATTCATAATCTGTTCCG 3214  
QY 3421 TTTTCTGATAGGAGCAGAAAGGCACTCTTCAATTTCTGTGAGGAGGCTTATAGAAAGCT 3480  
Db 3215 TTTTCTGATAGGAGCAGAAAGGCACTCTTCAATTTCTGTGAGGAGGCTTATAGAAAGCT 3274  
QY 3481 CAATTTTGCACAATTCATAGACCTTCAGAGGAGCTTCTTTAGGTTTGGATTTCTTCTT 3540  
Db 3275 CAATTTTGCACAATTCATAGACCTTCAGAGGAGCTTCTTTAGGTTTGGATTTCTTCTT 3334  
QY 3541 TAGGTTGTTCTTGGTGATCTTGGCTTGGCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3600  
Db 3335 TAGGTTGTTCTTGGTGATCTTGGCTTGGCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3394  
QY 3601 ACTTCATATCAGGTTTCTTCCACTGCTGCTCAACGCTCACACCGTACTTGGCAGCATCTAA 3660  
Db 3395 ACTTCATATCAGGTTTCTTCCACTGCTGCTCAACGCTCACACCGTACTTGGCAGCATCTAA 3454  
QY 3661 CTAAATGCAAAATAAAATAAGTCAGACATCTCCAGGCTATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3720  
Db 3455 CTAAATGCAAAATAAAATAAGTCAGACATCTCCAGGCTATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3514  
QY 3721 CTGCAAGTTCTATCAGCTTCTTCCCTTAATTTTGTAGCGTTCAACAAACTTCTGCTGCTCAATA 3780  
Db 3515 CTGCAAGTTCTATCAGCTTCTTCCCTTAATTTTGTAGCGTTCAACAAACTTCTGCTGCTCAATA 3574  
QY 3781 ACCGTTTGGTATAGAACCTTCTGAGCATTTCTTACGATCCCAAGGCTGCTTCCA 3840  
Db 3575 ACCGTTTGGTATAGAACCTTCTGAGCATTTCTTACGATCCCAAGGCTGCTTCCA 3634  
QY 3841 TGGCTCTAAGACCTTGTGATTGGCCAAACAGGAGTGGCTTCCAAGTGACAGAAACCAA 3900  
Db 3635 TGGCTCTAAGACCTTGTGATTGGCCAAACAGGAGTGGCTTCCAAGTGACAGAAACCAA 3694  
QY 3901 CACCTGTTTGTTCACACAAATTTCAAGCAGCTTCCATCACTCAATCCAAATTCGATACCCA 3960  
Db 3695 CACCTGTTTGTTCACACAAATTTCAAGCAGCTTCCATCACTCAATCCAAATTCGATACCCA 3754  
QY 3961 GCAACTTTTGTAGTTGCTCTCAGATGTAGCACCTTTATACCAAAACCGTGTGACGAGAGATT 4020  
Db 3755 GCAACTTTTGTAGTTGCTCTCAGATGTAGCACCTTTATACCAAAACCGTGTGACGAGAGATT 3814  
QY 4021 GGTAGCTCCAGTTTGTGCTTATAGCTTCGGAAATAGATTTTGGACAGTACACCA 4080  
Db 3815 GGTAGCTCCAGTTTGTGCTTATAGCTTCGGAAATAGATTTTGGACAGTACACCA 3874  
QY 4081 GSCCCAAACGAGTAATTAAGAGAGTCAGCCACCAAAAGTAGTGAATAGACCAATCGGGGGCT 4140  
Db 3875 GSCCCAAACGAGTAATTAAGAGAGTCAGCCACCAAAAGTAGTGAATAGACCAATCGGGGGCT 3934  
QY 4141 CAGTAGTCAAGAGACCCCAACAAATTTCACTGACAGGGAATTTTGTGACATCTTTCAGAAA 4200  
Db 3935 CAGTAGTCAAGAGACCCCAACAAATTTCACTGACAGGGAATTTTGTGACATCTTTCAGAAA 3994  
QY 4201 GTTCTGTTTCTAGTAGTCAATTTGCCGAGCATCAATAATGGGGATTATACCAAGAACCAACAG 4260

Db 3995 GTTCGTAATTCAGTAGTCAATTCGCCGAGCATCAATATGGGGATTATACGAGCAACAG 4054  
Qy 4261 TGGAAAGTCACATCTACCAACTTTCCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4320  
Db 4055 TGGAAAGTCACATCTACCAACTTTCCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4114  
Qy 4321 CATTAGTGAAACTTTTCAAAATCGCCAGTCGAGGAAGAAAGGCACACGCGATCTAGCAT 4380  
Db 4115 CATTAGTGAAACTTTTCAAAATCGCCAGTCGAGGAAGAAAGGCACACGCGATCTAGCAT 4174  
Qy 4381 TAGCGGGCAAGATGCAACTTTATCAACCGAGGTCTTATAGATAACCCCTAGCGCTGGGA 4440  
Db 4175 TAGCGGGCAAGATGCAACTTTATCAACCGAGGTCTTATAGATAACCCCTAGCGCTGGGA 4234  
Qy 4441 TCATCCCTTTGACAACTCTTTCTGCCAAATCTAGTCCCAAAATCACTTCATTGATACCAT 4500  
Db 4235 TCATCCCTTTGACAACTCTTTCTGCCAAATCTAGTCCCAAAATCACTTCATTGATACCAT 4294  
Qy 4501 TATTGTACAACTTGAGCAAGTTGTGATCAGCTCCTCAAAATGGTCTCTGTAAACGGATG 4560  
Db 4295 TATTGTACAACTTGAGCAAGTTGTGATCAGCTCCTCAAAATGGTCTCTGTAAACGGATG 4354  
Qy 4561 ACTCAACTTGACAACTTTAACTTGAAGCTCAGTCGATTCAGTGAACCTTGATCAGTTGTGCA 4620  
Db 4355 ACTCAACTTGACAACTTTAACTTGAAGCTCAGTCGATTCAGTGAACCTTGATCAGTTGTGCA 4414  
Qy 4621 GCTGTCAGCAGCATAGGAAACACGGCTTTCTACCAAACTCAAGGAATTAACAAC 4680  
Db 4415 GCTGTCAGCAGCATAGGAAACACGGCTTTCTACCAAACTCAAGGAATTAACAAC 4474  
Qy 4681 CTGCAACTTTCGCTATGACGTAGCAAGGAAATGTATCACTTTGAAAGTGGACAGTGAG 4740  
Db 4475 CTGCAACTTTCGCTATGACGTAGCAAGGAAATGTATCACTTTGAAAGTGGACAGTGAG 4534  
Qy 4741 TGTAGTCTTGAGAAATCTGAGCGGTATTTTATATATCAGTGAATCAGTCAACAGGAG 4800  
Db 4535 TGTAGTCTTGAGAAATCTGAGCGGTATTTTATATATCAGTGAATCAGTCAACAGGAG 4594  
Qy 4801 TCCTCTAGCGCGACGATCGTGCGCA 4828  
Db 4595 TCCTCTAGCGCGACGATCGTGCGCGACCTGACGGGGGGGGGGCGGTGAGTCTGC 4654  
Qy 4829 4828  
Db 4655 CTCGTGAAAGGTGTTGCTGACTATACAGGCGCTGAATCGCCCATCATCCAGCCAGA 4714  
Qy 4829 4828  
Db 4715 AAGTGAGGAGCCACCGTTGATGAGAGCTTTGTTGAGGTGGACAGTTGGTGATTGGA 4774  
Qy 4829 4828  
Db 4775 ACTTTTGCTTTGCCACGGAACGGTCTGCGTTGTGCGGAAGATGCGTGATCTGATCCTTCA 4834  
Qy 4829 4828  
Db 4835 ACTCAGAAAAGTTGATTTATTAACAAAGCGCGCTCCCGTCAAGTCACGTAATGCT 4894  
Qy 4829 4828  
Db 4895 CTGCCAGTTTACAAACCAATTAACCAATTCGTATTAGAAAACTCATCGACATCAAAATG 4954  
Qy 4829 4828  
Db 4955 AAACCTGCAATTTATCATATCAGGATTATCAATACCATATTTTGA AAAAGCCGTTTCTG 5014  
Qy 4829 4828  
Db 5015 TAATGAGGAGAAAACCTCACGAGGCAAGTCCATAGGATGCAAGATCCTCGTATCGGTC 5074  
Qy 4829 4828

Db 5075 TGGATTCCGACTCGTCCAAACATCAATACACCTATTAAATTTCCCTCTCTCAAAAATAAG 5134  
Qy 4829 4828  
Db 5135 GTTATCAAGTGAGAAATCACCATGAGTACGACTGAATCCCGGTGAGAAATGGCAAAAGCTT 5194  
Qy 4829 4828  
Db 5195 ATGCATTTCTTTCCAGACTTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACT 5254  
Qy 4829 4828  
Db 5255 CGCATCAACCAAAACCGTTTATTCAATTCGTGATTGCGCTGAGCGAGACGAAATACCGGATC 5314  
Qy 4829 4828  
Db 5315 GCTGTTAAAGGACAAATTAACAAACAGGATCGAATGCAACCGCGCAGGAAACACTGCCAG 5374  
Qy 4829 4828  
Db 5375 CGCATCAACATATTTTCACTCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTT 5434  
Qy 4829 4828  
Db 5435 CCGGGGATCGCAGTGGTGAGTAACCATGTCATCATCAGGATACGATTAATAATGTTGAT 5494  
Qy 4829 4828  
Db 5495 GGTCCGAGAGGATATAAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATC 5554  
Qy 4829 4828  
Db 5555 ATTGGCAACGCTACCTTTGCCATGTTTCAGAAAAAACAATCTGCGCGATCGGGCTTCCCAT 5614  
Qy 4829 4828  
Db 5615 CAATCGATAGATTGTCGACCTGATTGCCGACATTTATCGCGACATTCGCGGCCAATTTATACCAT 5674  
Qy 4829 4828  
Db 5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGCGCTCGAGCAAGAGTTTCCCGTTGAAT 5734  
Qy 4829 4828  
Db 5735 ATGGCTCATACACCCCTTGTTATTTACTGTTTATGTAAGCAGACAGTTTATTTGTTTCATGA 5794  
Qy 4829 4828  
Db 5795 TGATATATTTTATCTTTGTCATGTAACATCAGAGATTTTGAGACACAAACGTTGCTTTC 5854  
Qy 4829 4828  
Db 5855 CCCCCCCCCCTGTCAGGTTCGGCATCACCGCGCCACAGTGGCTTCTGGCGCCTATAT 4879  
Qy 4880 CGCCGACATCACCGATGGGAAAGATCGGGCTCGCCACTTTCGGGCTCATGAGCGCTTGT 4939  
Db 5915 CGCCGACATCACCGATGGGAAAGATCGGGCTCGCCACTTTCGGGCTCATGAGCGCTTGT 5974  
Qy 4940 CGCGTGGGTATGTTGGCAGGCCCCCGTGGCGGGGACTGTTGGCGCCATCTCTCTGCA 4999  
Db 5975 CGCGTGGGTATGTTGGCAGGCCCCCGTGGCGGGGACTGTTGGCGCCATCTCTCTGCA 6034  
Qy 5000 TGCACCATTTCTTGGCGGGCGGTGCTCAACGGCGCTCAACCTACTACTGGGTGCTTCT 5059  
Db 6035 TGCACCATTTCTTGGCGGGCGGTGCTCAACGGCGCTCAACCTACTACTGGGTGCTTCT 6094  
Qy 5060 AATGCAAGGATCGCATAAAGGAGAGCGTTCAGATATCTATGATGGAAGTATGGGAATGTT 5119  
Db 6095 AATGCAAGGATCGCATAAAGGAGAGCGTTCAGATATCTATGATGGAAGTATGGGAATGTT 6154  
Qy 5120 GATACCGCATTTCTTCAAGTGTCTTGGGTCTCTATCAGATTCCTCAACCTAAAGCAAC 5179  
Db 6155 GATACCGCATTTCTTCAAGTGTCTTGGGTCTCTATCAGATTCCTCAACCTAAAGCAAC 6214

QY 5180 CGAGAGGAGAGATTTCAATGTAATAATTTCTCTGACTTTTGGTCAATCAAGTACAGCTGAACTG 5239  
Db 6215 CGAGAGGAGAGATTTCAATGTAATAATTTCTCTGACTTTTGGTCAATCAAGTACAGCTGAACTG 6274  
QY 5240 TGAGACTATCTCGGTATATGACAGAGAAATGTCCTCTCTGAGACAGAGTAAATGAAGTCCC 5299  
Db 6275 TGAGACTATCTCGGTATATGACAGAGAAATGTCCTCTCTGAGACAGAGTAAATGAAGTCCC 6334  
QY 5300 ACCAATAAGAGAAATCCTTGTATACAGAAACAAATCTCTCTCTGAGACAGAGTAAATGAAGTCCC 5359  
Db 6335 ACCAATAAGAGAAATCCTTGTATACAGAAACAAATCTCTCTCTGAGACAGAGTAAATGAAGTCCC 6394  
QY 5360 TTGAACATAAATGTAAGTATGATGTCGGGTAGGAGTAAATGAAGTAAATGCTTACC 5419  
Db 6395 TTGAACATAAATGTAAGTATGATGTCGGGTAGGAGTAAATGAAGTAAATGCTTACC 6454  
QY 5420 TTCTGGACCTTCAAGAGGTATGAGGTATGTAAGTATGATGTCGATGCCAACTTCAAGTGAACAAC 5479  
Db 6455 TTCTGGACCTTCAAGAGGTATGAGGTATGTAAGTATGATGTCGATGCCAACTTCAAGTGAACAAC 6514  
QY 5480 GTTCTATTTCTGTTCAAAACCAATCCGAATCCAGAGAAATCAAGTGTGTTGTTCTACTATT 5539  
Db 6515 GTTCTATTTCTGTTCAAAACCAATCCGAATCCAGAGAAATCAAGTGTGTTGTTCTACTATT 6574  
QY 5540 GATCCAGGCGAGTGGGTCTTGAACCTGCAATAGTGTGCTGTTGTTGAGGTCTATCTT 5599  
Db 6575 GATCCAGGCGAGTGGGTCTTGAACCTGCAATAGTGTGCTGTTGTTGAGGTCTATCTT 6634  
QY 5600 TGTAAGATAAATCTAGTCTTTGATTAATAATCTTGAAGAGGAGGAGGAGTAATAC 5659  
Db 6635 TGTAAGATAAATCTAGTCTTTGATTAATAATCTTGAAGAGGAGGAGGAGTAATAC 6694  
QY 5660 CAATCTAAACCTCTTTAAACCTTAAAGGAGCAAGTATGTCGCTGCTGATTAATAACCCC 5719  
Db 6695 CAATCTAAACCTCTTTAAACCTTAAAGGAGCAAGTATGTCGCTGCTGATTAATAACCCC 6754  
QY 5720 AAATCAGCTCTAGTCTGATCCTCATCAACTTGAAGGAGCAATCTTGTGTTTGAAGAAAT 5779  
Db 6755 AAATCAGCTCTAGTCTGATCCTCATCAACTTGAAGGAGCAATCTTGTGTTTGAAGAAAT 6814  
QY 5780 TTGGGAGATCGATATCGAGAAAGGAGTACGCTGATTTTAAAGCGTAAATTTATCTCAA 5839  
Db 6815 TTGGGAGATCGATATCGAGAAAGGAGTACGCTGATTTTAAAGCGTAAATTTATCTCAA 6874  
QY 5840 GATCTCTGCTCGCGCTTTTGGGTGATGACGGTGAACCACTCTGACATGACAGCTCCCG 5899  
Db 6875 GATCTCTGCTCGCGCTTTTGGGTGATGACGGTGAACCACTCTGACATGACAGCTCCCG 6934  
QY 5900 GAGACGCTCAGCTGTTGTTGTAAGCGGATGTCGGGAGGAGAGCAAGCCGTCAGGCGCG 5959  
Db 6935 GAGACGCTCAGCTGTTGTTGTAAGCGGATGTCGGGAGGAGAGCAAGCCGTCAGGCGCG 6994  
QY 5960 TCAGCGGTGTTGCGGGTTCGGGCGCAGCAATGACCCAGTCAAGTACCGATAGCGGA 6019  
Db 6995 TCAGCGGTGTTGCGGGTTCGGGCGCAGCAATGACCCAGTCAAGTACCGATAGCGGA 7054  
QY 6020 GTGTATCTGCTTAACTATGCGGCATCAGAGCAAGTGTACTGAGAGTGCACATATGC 6079  
Db 7055 GTGTATCTGCTTAACTATGCGGCATCAGAGCAAGTGTACTGAGAGTGCACATATGC 7114  
QY 6080 GGTGTGAATACCGCAGATGCTTAAGAGAGAAATACCGCATCAGCGCTCTCCGCTT 6139  
Db 7115 GGTGTGAATACCGCAGATGCTTAAGAGAGAAATACCGCATCAGCGCTCTCCGCTT 7174  
QY 6140 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 6199  
Db 7175 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCACT 7234  
QY 6200 CAAGGCGGTAAATCGGTATTCACAGATCAGGGGATACCGAGGAGAAACATGTGAG 6259  
Db 7235 CAAGGCGGTAAATCGGTATTCACAGATCAGGGGATACCGAGGAGAAACATGTGAG 7294

QY 6260 CAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCGCGTGTCTGGCGTTTTCCTAATA 6319  
Db 7295 CAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCGCGTGTCTGGCGTTTTCCTAATA 7354  
QY 6320 GGCTCGCGCCCTCTGACGAGCATCACAAAAATCGACCTCAAGTCAAGGTGCGGAAACC 6379  
Db 7355 GGCTCGCGCCCTCTGACGAGCATCACAAAAATCGACCTCAAGTCAAGGTGCGGAAACC 7414  
QY 6380 CGACAGGACTATAAGATACAGAGCTTTCCCTCTGGAAGTCCCTCGTGGCTCTCTCTG 6439  
Db 7415 CGACAGGACTATAAGATACAGAGCTTTCCCTCTGGAAGTCCCTCGTGGCTCTCTCTG 7474  
QY 6440 TTCCGACCTCGCGCTTACCGGATACCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 6499  
Db 7475 TTCCGACCTCGCGCTTACCGGATACCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 7534  
QY 6500 TTCTCATAGCTCAGCTGATGATCTCTCAGTTCGGTGTAGTCTGTTTCGCTCCAGCTGG 6559  
Db 7535 TTCTCATAGCTCAGCTGATGATCTCTCAGTTCGGTGTAGTCTGTTTCGCTCCAGCTGG 7594  
QY 6560 GCTGTGTGCAAGAACCCCGCTTCAAGCCGACCGCTTATCCGGTAACTATCTCTCT 6619  
Db 7595 GCTGTGTGCAAGAACCCCGCTTCAAGCCGACCGCTTATCCGGTAACTATCTCTCT 7654  
QY 6620 TTGAGTCAACCCCGTAAAGACAGCACTTATCGCACTGGCAGAGCCACTGGTAAACAGGA 6679  
Db 7655 TTGAGTCAACCCCGTAAAGACAGCACTTATCGCACTGGCAGAGCCACTGGTAAACAGGA 7714  
QY 6680 TTGAGCAGGCGAGTATGATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACATACG 6739  
Db 7715 TTGAGCAGGCGAGTATGATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACATACG 7774  
QY 6740 GCTACACTAGAAAGACAGTATTTGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6799  
Db 7775 GCTACACTAGAAAGACAGTATTTGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7834  
QY 6800 AAAGAGTGGTGTAGTCTTGTATCCGCAACCAACCCGCTGCTGCTGCTGCTGCTGCTGCT 6859  
Db 7835 AAAGAGTGGTGTAGTCTTGTATCCGCAACCAACCCGCTGCTGCTGCTGCTGCTGCTGCT 7894  
QY 6860 TTGCAAGCAGCAGATTAAGCGCGCAAGAAAGAGTCTCAAGAAAGTCTTGTGATCTTTT 6919  
Db 7895 TTGCAAGCAGCAGATTAAGCGCGCAAGAAAGAGTCTCAAGAAAGTCTTGTGATCTTTT 7954  
QY 6920 CTAAGGCTGTAGCGCTCAGTGGAGCAAGAACTCAGCTTAAAGGATTTTGGTCAATGAGAT 6979  
Db 7955 CTAAGGCTGTAGCGCTCAGTGGAGCAAGAACTCAGCTTAAAGGATTTTGGTCAATGAGAT 8014  
QY 6980 TATCAAAAGGATCTTCACTTGTATTTTAAATTTAAATTTAAATTTAAATTTAAATCT 7039  
Db 8015 TATCAAAAGGATCTTCACTTGTATTTTAAATTTAAATTTAAATTTAAATTTAAATCT 8074  
QY 7040 AAGATATATAGTAAACTTGGTCTGACATTAACAAATGCTTAATCAGTGAAGCACTTA 7099  
Db 8075 AAGATATATAGTAAACTTGGTCTGACATTAACAAATGCTTAATCAGTGAAGCACTTA 8134  
QY 7100 TCTCAGCGATCTGCTATTTCTGTTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7159  
Db 8135 TCTCAGCGATCTGCTATTTCTGTTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8194  
QY 7160 CTAAGATACGGAGGCTTACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7219  
Db 8195 CTAAGATACGGAGGCTTACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8254  
QY 7220 GCTCAGCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGAGGCGGAGCGAGAA 7279  
Db 8255 GCTCAGCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGAGGCGGAGCGAGAA 8314  
QY 7280 GTGTCTCTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGAG 7339  
Db 8315 GTGTCTCTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGAG 8374  
QY 7340 TAAGTAGTCTGCGCAGTAAATAGTTTGGCAACGTTGTTGCCATTGCTGCGGATCTGTTG 7399

8375	TAAAGTAGTTGCCCAGGTAAAGTGTTCGCAACGTTGTTCGCAATGTGTGCAGGCATCGTGG	8434
7400	TGTCACGCTCGTGGTTGGTATGGCTTCATTACAGTCCGGTTCGCCAACGATCAAGCGGAG	7459
8435	TGTCACGCTCGTGGTTGGTATGGCTTCATTACAGTCCGGTTCGCCAACGATCAAGCGGAG	8494
7460	TTACATGATCCCCCATGTTCTGTCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCGTTTG	7519
8495	TTACATGATCCCCCATGTTCTGTCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCGTTTG	8554
7520	TCAGAGTAGTTGGCCGCGAGTGTATCACTCATGTTTATGGCAGGACTGCATAATTCTC	7579
8555	TCAGAGTAGTTGGCCGCGAGTGTATCACTCATGTTTATGGCAGGACTGCATAATTCTC	8614
7580	TTACTGTCAATCCGATCCGTAAGATGCTTTTCTGTGACTGTTGAGTACTCAACCAAGTCAT	7639
8615	TTACTGTCAATCCGATCCGTAAGATGCTTTTCTGTGACTGTTGAGTACTCAACCAAGTCAT	8674
7640	TCGTAGAAATAGTATGTCGGCGGACCGAGTTGCTCTTCGCCGGCGTCAACACGGGATAATA	7699
8675	TCGTAGAAATAGTATGTCGGCGGACCGAGTTGCTCTTCGCCGGCGTCAACACGGGATAATA	8734
7700	CCGGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGAAAAAGTTTCTTCGGGGCGAA	7759
8735	CCGGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGAAAAAGTTTCTTCGGGGCGAA	8794
7760	AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCTGTGCACCCA	7819
8795	AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCTGTGCACCCA	8854
7820	ACTGATCTTCAGCATCTTTTACTTTCCACGAGCGTTTCTGGGTGAGCAAAACACGGAAGGC	7879
8855	ACTGATCTTCAGCATCTTTTACTTTCCACGAGCGTTTCTGGGTGAGCAAAACACGGAAGGC	8914
7880	AAAATGCGCAAAAAGGGGAATGAAGGCGACACGAAATGTGAATACTCATACTCTTCC	7939
8915	AAAATGCGCAAAAAGGGGAATGAAGGCGACACGAAATGTGAATACTCATACTCTTCC	8974
7940	TTTTTCAATATATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTG	7999
8975	TTTTTCAATATATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTG	9034
8000	AATGTATTTAGAAAAATAACAAATAGGGTTCCGGCGACATTTCCCGCAAAAGTGCCAC	8059
9035	AATGTATTTAGAAAAATAACAAATAGGGTTCCGGCGACATTTCCCGCAAAAGTGCCAC	9094
8060	CTGACGCTCAAGAAACCATTTATTATCATGACATTAACCTATAAAAAATAGCGTATCACGA	8119
9095	CTGACGCTCAAGAAACCATTTATTATCATGACATTAACCTATAAAAAATAGCGTATCACGA	9154
8120	GGCCCTTCGTTTCAGAAATTAATTCTCATGTTTGACGCTTATCATCATAGCTGAC	8179
9155	GGCCCTTCGTTTCAGAAATTAATTCTCATGTTTGACGCTTATCATCATAGCTGAC	9214
8180	TCATGTTGGTATTGTGAAATAGACCGCAGATCGGGAACTGTGAAAAATAACAGTTATTATT	8239
9215	TCATGTTGGTATTGTGAAATAGACCGCAGATCGGGAACTGTGAAAAATAACAGTTATTATT	9274
8240	CG 8241	
9275	CG 9276	

#### RESULT 4

RESULT 4  
 US-10-267-219-13  
 ; Sequence 13, Application US/10267219  
 ; Publication No. US20030143219A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Madison, Edwin  
 ; APPLICANT: Yeh, Jiunn-Chern  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING A TRANSMEMBRANE SERINE PROTEASE 2  
 ; TITLE OF INVENTION: ENCODED POLYPEPTIDES AND METHODS BASED THEREON

```

; FILE REFERENCE: 24745-1621
; CURRENT APPLICATION NUMBER: US/10/267,219
; CURRENT FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: 60/328,530
; PRIOR FILING DATE: 09-OCT-2001
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
US-10-267-219-13

```

Query Match	79.6%;	Score 6561.6;	DB 15;	Length 9276;
Best Local Similarity	84.7%;	Pred. No. 0;		
Matches 8031;	Conservative	0;	Mismatches	4;
				Indels 1447;
				Gaps
				2;
Qy	1	AGATCTTAACATCAAAGACGAAAGGTTGAATGAAACCTTTTGGCATCCGACATCCACAG	60	
Db	1	AGATCTTAACATCAAAGACGAAAGGTTGAATGAAACCTTTTGGCATCCGACATCCACAG	60	
Qy	61	GTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACATAGCAGCAGACCGT	120	
Db	61	GTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACATAGCAGCAGACCGT	120	
Qy	121	TGCAAGCAGGAGCCTCCACTCTCTCTCTCTCAACACCACTTTTGCCATCGAAAAACC	180	
Db	121	TGCAAGCAGGAGCCTCCACTCTCTCTCTCTCAACACCACTTTTGCCATCGAAAAACC	180	
Qy	181	AGCCCATGTTATTGGGCTTGATTGGAGCTCGCTCATTCCTCAATTCCTTCTATTAGGCTACTA	240	
Db	181	AGCCCATGTTATTGGGCTTGATTGGAGCTCGCTCATTCCTCAATTCCTTCTATTAGGCTACTA	240	
Qy	241	ACACCATGACTTTATTAGCCTGTCTATCTCTGGGCCGCCCTGGCGAGGTTTCATGTTTGTGTTA	300	
Db	241	ACACCATGACTTTATTAGCCTGTCTATCTCTGGGCCGCCCTGGCGAGGTTTCATGTTTGTGTTA	300	
Qy	301	TTTCCGAATGCAACCAAGCTCCGATTACACCCGAAACATCACTCCAGATGAGGGCTTTCGT	360	
Db	301	TTTCCGAATGCAACCAAGCTCCGATTACACCCGAAACATCACTCCAGATGAGGGCTTTCGT	360	
Qy	361	AGTGTGGGTGCAAAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAAGCCT	420	
Db	361	AGTGTGGGTGCAAAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAAGCCT	420	
Qy	421	GTCTTGGAACCTTAATGACAAAGCGTGATCTCATCCAGATGCAACTAAGTTTGGTTCG	480	
Db	421	GTCTTGGAACCTTAATGACAAAGCGTGATCTCATCCAGATGCAACTAAGTTTGGTTCG	480	
Qy	481	TTGAAATGCTAAACGGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCCATACCGTTTGT	540	
Db	481	TTGAAATGCTAAACGGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCCATACCGTTTGT	540	
Qy	541	CTTGTTTGGTATTGATTGACGAATGCTCAAAATTAATCTCATTAATGCTTAGCGCAGTCT	600	
Db	541	CTTGTTTGGTATTGATTGACGAATGCTCAAAATTAATCTCATTAATGCTTAGCGCAGTCT	600	
Qy	601	CTCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAACGCAATGGGGAACACACCGCT	660	
Db	601	CTCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAACGCAATGGGGAACACACCGCT	660	
Qy	661	TTTTGGAGATTATGCAATGCTCCCAATTTGATGCTTCCAGATTCCTGGTGGGAATACT	720	
Db	661	TTTTGGAGATTATGCAATGCTCCCAATTTGATGCTTCCAGATTCCTGGTGGGAATACT	720	
Qy	721	GCTGATAGCCTAAGTTTCATGATCAAAATTTAAGTTCTTAACCCCTACTTTGACGCAAT	780	
Db	721	GCTGATAGCCTAAGTTTCATGATCAAAATTTAAGTTCTTAACCCCTACTTTGACGCAAT	780	
Qy	781	ATATAAACAGAGGAAGCTCCCTGTCTTAAACCTTTTTTTTATCATCATTAATTAAGCTT	840	
Db	781	ATATAAACAGAGGAAGCTCCCTGTCTTAAACCTTTTTTTTATCATCATTAATTAAGCTT	840	

841 ACTTTCATAATGGACGTGGTTCCAAATGACAAGCTTTTGATTTTAAACGACTTTTAAACGA 900  
841 ACTTTCATAATGGACGTGGTTCCAAATGACAAGCTTTTGATTTTAAACGACTTTTAAACGA 900  
901 CAACCTTGAGAGATCAAAAAACAATAATATTCGAAGGATCCAAACGATGAGATTTCCCT 960  
901 CAACCTTGAGAGATCAAAAAACAATAATATTCGAAGGATCCAAACGATGAGATTTCCCT 960  
961 TCAATTTTACTGCAAGTTTATTCGACGATCCCTCCGATTTAGCTGCTCCAGTCAACACT 1020  
961 TCAATTTTACTGCAAGTTTATTCGACGATCCCTCCGATTTAGCTGCTCCAGTCAACACT 1020  
1021 ACAACAGAAGATGAACCGGCAAAATCCCGCTGAAGCTGTCATCGGTTTACTCAGATTTA 1080  
1021 ACAACAGAAGATGAACCGGCAAAATCCCGCTGAAGCTGTCATCGGTTTACTCAGATTTA 1080  
1081 GAAGGGGATTTGATGTTGCTGTTTGCATTTTCCACAGCACAATAACGGGTTATTG 1140  
1081 GAAGGGGATTTGATGTTGCTGTTTGCATTTTCCACAGCACAATAACGGGTTATTG 1140  
1141 TTTATAAATACTACTATTGCCAGCATTCGCTGCTAAAGAAAGGGGTATCTCTCGAGAAA 1200  
1141 TTTATAAATACTACTATTGCCAGCATTCGCTGCTAAAGAAAGGGGTATCTCTCGAGAAA 1200  
1201 AGAGAGGCTGAAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGGAAACAATGATCATG 1260  
1201 AGAGAGGCTGAAGCTTA----- 1217  
1261 GCCCCCCGGAGAGGATAAATGTGGCTTCCCGGTGTCACCGCCAGCAGTGCACGGAG 1320  
1218 ----- 1217  
1321 AGAGGTTGCTGTTTGTGATGACAGTGTCCGGGATTCGCGTGGTCTCCACCCCATGGCC 1380  
1218 ----- 1217  
1381 ATCGAGACACTCAAGAGAGAAGATGTCCCTTCTAACTAGTGGGTGAGAAATTCCTTAGGG 1440  
1218 -----CGTAGAATTCCTTAGGG 1234  
1441 CGCCGCGAAATTAATTCGCTTAGACATGACTGTTCTCAGTCAAGTGGGCACTTACG 1500  
1235 CGCCGCGAAATTAATTCGCTTAGACATGACTGTTCTCAGTCAAGTGGGCACTTACG 1294  
1501 AGAGACCGGCTTCTGCTAGATTTAAATCAAGAGATGTCAGATGCCATTTGCTGAGAG 1560  
1295 AGAGACCGGCTTCTGCTAGATTTAAATCAAGAGATGTCAGATGCCATTTGCTGAGAG 1354  
1561 ATGACGGCTTCATTTTTCATATCTTTTATTTTGTAACTATATAGTATAGGATTTTTTTT 1620  
1355 ATGACGGCTTCATTTTTCATATCTTTTATTTTGTAACTATATAGTATAGGATTTTTTTT 1414  
1621 GTCATTTTGTCTTCTCGTACGAGCTTGTCTGATCAGCCCTATCTCGCAGCTGATGAA 1680  
1415 GTCATTTTGTCTTCTCGTACGAGCTTGTCTGATCAGCCCTATCTCGCAGCTGATGAA 1474  
1681 TATCTTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTTGTTATTTCCC 1740  
1475 TATCTTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTTGTTATTTCCC 1534  
1741 ACTCCTCTTCAGAGPACAGAGATTAAGTGAAGAGTTGTTGTTGCAAGCTTATTCGATAA 1800  
1535 ACTCCTCTTCAGAGPACAGAGATTAAGTGAAGAGTTGTTGTTGCAAGCTTATTCGATAA 1594  
1801 GCTTTAATTCGGTATGTTTATCACAGTTTAAATGCTTAACGAGTCCAGCCCGCTGATGAA 1860  
1595 GCTTTAATTCGGTATGTTTATCACAGTTTAAATGCTTAACGAGTCCAGCCCGCTGATGAA 1654  
1861 ATCTAACAAATCGCTCATCTGTCATCTCGGACCGCTGACCCCTGGATGCTGTAGGCATPAG 1920  
1655 ATCTAACAAATCGCTCATCTGTCATCTCGGACCGCTGACCCCTGGATGCTGTAGGCATPAG 1714  
1921 CTGTTATGCGGTACTGCGCGGGCTTCTCGGGGATATCGTCCATTCGACAGCATCGC 1980

1715 CTTGGTTATGCCGGTACTGCGGGGCTCTTGGGATATCGTCCATTCGACAGCATCGC 1774  
1981 CAGTCACACTATGCGGTGCTAGCGCTATATGCGTTGATGCAATTTCTATCGCACCCGT 2040  
1775 CAGTCACACTATGCGGTGCTAGCGCTATATGCGTTGATGCAATTTCTATCGCACCCGT 1834  
2041 TCTCGAGACACTGTCGACCGCTTTGGCGCGCCGACGCTCTGCTCGCTTCTGCTACTTGG 2100  
1835 TCTCGAGACACTGTCGACCGCTTTGGCGCGCCGACGCTCTGCTCGCTTCTGCTACTTGG 1894  
2101 AGCCACTATCGACTAGCGGATCATGCGGACACACACCGCTCTGCTGATCTATCGAATCTA 2160  
1895 AGCCACTATCGACTAGCGGATCATGCGGACACACACCGCTCTGCTGATCTATCGAATCTA 1954  
2161 AATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATACGAACCTTAA 2220  
1955 AATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATACGAACCTTAA 2014  
2221 CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGAGCCGAGATCCATCTGCTTGGCCAA 2280  
2015 CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGAGCCGAGATCCATCTGCTTGGCCAA 2074  
2281 TATGTTTCACTCCCTCAGGATTTAGTCTTGTGAAGTGAATGAACTTCTGGAAGTTGAG 2340  
2075 TATGTTTCACTCCCTCAGGATTTAGTCTTGTGAAGTGAATGAACTTCTGGAAGTTGAG 2134  
2341 TGTAACTCCGCTGTTATTGACGGGATATCCGTAGTTGGCAAGTGTGTTGTTGTTACGG 2400  
2135 TGTAACTCCGCTGTTATTGACGGGATATCCGTAGTTGGCAAGTGTGTTGTTGTTACGG 2194  
2401 AGAGTAACTCCCAACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGCTGTT 2460  
2195 AGAGTAACTCCCAACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGCTGTT 2254  
2461 GTACTTGTATCAATAGAAGACATTTCTCGATTTGACAGATCAAGTGTTCAGAGCGT 2520  
2255 GTACTTGTATCAATAGAAGACATTTCTCGATTTGACAGATCAAGTGTTCAGAGCGT 2314  
2521 ACTGATTTGACATTTCCAAAGCTCTCTGTTAGTTGCAACCGATAGGTTGTAGAGTGTG 2580  
2315 ACTGATTTGACATTTCCAAAGCTCTCTGTTAGTTGCAACCGATAGGTTGTAGAGTGTG 2374  
2581 CAATACACTTGGTACAAATTTCAACCTTTGCACTGACAGCTTGGTTGTGAACAGCAT 2640  
2375 CAATACACTTGGTACAAATTTCAACCTTTGCACTGACAGCTTGGTTGTGAACAGCAT 2434  
2641 CTTCAATTTGCGCAAGCTCCTTGTCTGATATCGACAGCCAAACAGATACCTGGGAAT 2700  
2435 CTTCAATTTGCGCAAGCTCCTTGTCTGATATCGACAGCCAAACAGATACCTGGGAAT 2494  
2701 CAATACCAATTTGACATTTGAGACAGAGTCTGAGGCAACGAAATCTCGATCAGCGTATT 2760  
2495 CAATACCAATTTGACATTTGAGACAGAGTCTGAGGCAACGAAATCTCGATCAGCGTATT 2554  
2761 TATCAGCAATAACTAGAACTTCTGAGAGCCCGCAGCAGGATGTCATACTACAGAGGCTG 2820  
2555 TATCAGCAATAACTAGAACTTCTGAGAGCCCGCAGCAGGATGTCATACTACAGAGGCTG 2614  
2821 ATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAAACGATGTTTCTCTCGACCAATA 2880  
2615 ATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAAACGATGTTTCTCTCGACCAATA 2674  
2881 TTTTGTACACACTTAGGAAACAGTTTCTGTTCCGTAAAGCCATAGCCTACTGCTGGGCGC 2940  
2675 TTTTGTACACACTTAGGAAACAGTTTCTGTTCCGTAAAGCCATAGCCTACTGCTGGGCGC 2734  
2941 CTCCTGCTAGCAGATACACTTTAGCACCAACCTTGTGGGCAACGTAGTACTTCTGGGG 3000  
2735 CTCCTGCTAGCAGATACACTTTAGCACCAACCTTGTGGGCAACGTAGTACTTCTGGGG 2794  
3001 TAAGGCTACCATCTTCTTAGTGGAGATGCAAAAAACAATTTCTTTGCAACAGCACTT 3060

Db 2795 TAAGGGTACCATCCTCTTAGTGGAGATGCAAAAACAAATTTCTTTGCCAACCGAACTT 2854  
Qy 3061 TGGCAGGAACACCCAGCATCAGGGAAGTGGAGGCAGAAATGGCGTTTCCACGAGGAATAT 3120  
Db 2855 TGGCAGGAACACCCAGCATCAGGGAAGTGGAGGCAGAAATGGCGTTTCCACGAGGAATAT 2914  
Qy 3121 AGAGGCCAATCTTCTCAATAGGTTCTTCAAAACGAGAGCAGACTACACGAGGCAAGTCT 3180  
Db 2915 AGAGGCCAATCTTCTCAATAGGTTCTTCAAAACGAGAGCAGACTACACGAGGCAAGTCT 2974  
Qy 3181 CAATCTGCAACGCTCTCCGTTAGTTAGCTTCATGGAATTTCCCTGACCTTATCTATAGAGA 3240  
Db 2975 CAATCTGCAACGCTCTCCGTTAGTTAGCTTCATGGAATTTCCCTGACCTTATCTATAGAGA 3034  
Qy 3241 GATCAATGGCTCTCTTAACGTTAATCTGGCAATTCGTAAGTTCTCTGGGAAGAGGCTT 3300  
Db 3035 GATCAATGGCTCTCTTAACGTTAATCTGGCAATTCGTAAGTTCTCTGGGAAGAGGCTT 3094  
Qy 3301 CTAACACAGGTGCTCTCAAGCGACTCCATCAAACTTTGGGAGTTAGTTCTTAAAGGGCTT 3360  
Db 3095 CTAACACAGGTGCTCTCAAGCGACTCCATCAAACTTTGGGAGTTAGTTCTTAAAGGGCTT 3154  
Qy 3361 TGTCAACATTTGACGAACATTTGCGACAATTTGGTTTGACATAATTCATATCTGTTCCG 3420  
Db 3155 TGTCAACATTTGACGAACATTTGCGACAATTTGGTTTGACATAATTCATATCTGTTCCG 3214  
Qy 3421 TTTTCTGATAGGACGAGGAGGCTCTCAATTTCTTGTGAGGAGGCTTGAACGCT 3480  
Db 3215 TTTTCTGATAGGACGAGGAGGCTCTCAATTTCTTGTGAGGAGGCTTGAACGCT 3274  
Qy 3481 CAATTTTGCAAAATCAATAGCAGCTTCAGAAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3540  
Db 3275 CAATTTTGCAAAATCAATAGCAGCTTCAGAAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3334  
Qy 3541 TAGTTGTTCTTGGTGTATCTTGGCTTGGCATCTCTCTTCTTCTAGTACCTTTAGGG 3600  
Db 3335 TAGTTGTTCTTGGTGTATCTTGGCTTGGCATCTCTCTTCTTCTAGTACCTTTAGGG 3394  
Qy 3601 ACTTCATATCAGGTTTCTCTCCACCTCGTCCAACTGTCACACGCTACTTGGCAATCTAA 3660  
Db 3395 ACTTCATATCAGGTTTCTCTCCACCTCGTCCAACTGTCACACGCTACTTGGCAATCTAA 3454  
Qy 3661 CTAATGCAAAATAAAATAAGTCAGCAGCATTCGCCAGGCTATATCTTCTGATTTAGCTT 3720  
Db 3455 CTAATGCAAAATAAAATAAGTCAGCAGCATTCGCCAGGCTATATCTTCTGATTTAGCTT 3514  
Qy 3721 CTGCAAGTTCAATCAGCTTCTCCCTTAATTTTAGCGTTCAACAACTTCGTCGTCAAATA 3780  
Db 3515 CTGCAAGTTCAATCAGCTTCTCCCTTAATTTTAGCGTTCAACAACTTCGTCGTCAAATA 3574  
Qy 3781 ACCGTTTGGTATAAGAACCTTCTCGAGCAATTCCTTTACGATCCCAAGGTGGCTTCCA 3840  
Db 3575 ACCGTTTGGTATAAGAACCTTCTCGAGCAATTCCTTTACGATCCCAAGGTGGCTTCCA 3634  
Qy 3841 TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAGTGACAGAAACCAA 3900  
Db 3635 TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAGTGACAGAAACCAA 3694  
Qy 3901 CACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3960  
Db 3695 CACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3754  
Qy 3961 GCACTTTTGTGTTGCTCCAGATAGCACTTTATACCAAAACCGGTGACGAGATT 4020  
Db 3755 GCACTTTTGTGTTGCTCCAGATAGCACTTTATACCAAAACCGGTGACGAGATT 3814  
Qy 4021 GGTAGCTCCAGTTTGTCTCTTATAGCTTCGGGAATAGACTTTTGGAGAGGTACACCA 4080  
Db 3815 GGTAGCTCCAGTTTGTCTCTTATAGCTTCGGGAATAGACTTTTGGAGAGGTACACCA 3874  
Qy 4081 GGCCCAACGAGTAATTAGAAGAGTCAGCCACCAAAAGTAGTGAATAGACCTCGGGCGGT 4140  
Db 3875 GGCCCAACGAGTAATTAGAAGAGTCAGCCACCAAAAGTAGTGAATAGACCTCGGGCGGT 3934

Qy 4141 CAGTAGTCAAGACGCCCAACAAAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAA 4200  
Db 3935 CAGTAGTCAAGACGCCCAACAAAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAA 3994  
Qy 4201 GTTCGTATTTCAGTAGTCAATTCGCCGAGCATCAATTAATGGGATTTATACAGAAAGCAACAG 4260  
Db 3995 GTTCGTATTTCAGTAGTCAATTCGCCGAGCATCAATTAATGGGATTTATACAGAAAGCAACAG 4054  
Qy 4261 TGAAGTGCATCTCTACCAACTTTTGGCTCTCAGAAAAAGCATATAACAGTTCTTACTACCGC 4320  
Db 4055 TGAAGTGCATCTCTACCAACTTTTGGCTCTCAGAAAAAGCATATAACAGTTCTTACTACCGC 4114  
Qy 4321 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCATCTAGCAT 4380  
Db 4115 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCATCTAGCAT 4174  
Qy 4381 TAGCGGGCAAGAGTCAACTTTTATCAACAGGGTCTCTATAGATAACCCCTAGCGCTGGGA 4440  
Db 4175 TAGCGGGCAAGAGTCAACTTTTATCAACAGGGTCTCTATAGATAACCCCTAGCGCTGGGA 4234  
Qy 4441 TCATCCTTTGGCAACTCTTTCTGCAAAATCTAGSTCCAAATCACTTCATTGATACCAT 4500  
Db 4235 TCATCCTTTGGCAACTCTTTCTGCAAAATCTAGSTCCAAATCACTTCATTGATACCAT 4294  
Qy 4501 TATTGTACAACTTGAGCAAGTTTGTGATCAGCTCCTCAAAATTTGGTCTCTGTAAACGATG 4560  
Db 4295 TATTGTACAACTTGAGCAAGTTTGTGATCAGCTCCTCAAAATTTGGTCTCTGTAAACGATG 4354  
Qy 4561 ACTCAACTTTGCACTTAATCTGAAGCTCAGTCGATTTAGTGAACTTTGATCAGGTTGTGCA 4620  
Db 4355 ACTCAACTTTGCACTTAATCTGAAGCTCAGTCGATTTAGTGAACTTTGATCAGGTTGTGCA 4414  
Qy 4621 GCTGGTTCAGCAGCATAGGGAACACGGCTTTTCTTACCAAACTCAAGGAATTTATCAAACT 4680  
Db 4415 GCTGGTTCAGCAGCATAGGGAACACGGCTTTTCTTACCAAACTCAAGGAATTTATCAAACT 4474  
Qy 4681 CTGCAACACTTCCGATGCAAGTAGCAGTAGCAAGGAAATGTCTACTTGAAGTCGACAGTGAG 4740  
Db 4475 CTGCAACACTTCCGATGCAAGTAGCAGTAGCAAGGAAATGTCTACTTGAAGTCGACAGTGAG 4534  
Qy 4741 TGTAGTCTTGAGAAATTTCTGAAGCCGTATTTTATATCAGTGAGTCAGTCATCAGGAGA 4800  
Db 4535 TGTAGTCTTGAGAAATTTCTGAAGCCGTATTTTATATCAGTGAGTCAGTCATCAGGAGA 4594  
Qy 4801 TCCTCTACGCCGAGCGCATCGTGGCCGA----- 4828  
Db 4595 TCCTCTACGCCGAGCGCATCGTGGCCGA----- 4654  
Qy 4829 ----- 4828  
Db 4655 CTCTGAGAGAGGTGTTGCTGACTCATACAGGCCCTGAATCGCCCATCATCCAGCCAGA 4714  
Qy 4829 ----- 4828  
Db 4715 AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGTGGACCAGTTGGTGATTGCA 4774  
Qy 4829 ----- 4828  
Db 4775 ACTTTTGTCTTCCACGGAACGGTCTGCTGTTCTGCGGAAGATGCGTGATCTCATCTTCA 4834  
Qy 4829 ----- 4828  
Db 4835 ACTCAGAAAAAGTTCGATTTATTTCACAAAGCCGCCCTCCCGTCAAGTCAGCGTAATGCT 4894  
Qy 4829 ----- 4828  
Db 4895 CTGCCAGTGTACAAACCAATTAACCAATTTCTGATTAGAAAAAACTCATCGAGCATCAAAATG 4954  
Qy 4829 ----- 4828  
Db 4955 AAACGTCAATTTTATTATATCAGGATTTATCAATACCATATTTTGGAAAAAGCCGTTCTG 5014



QY 4829 ----- 4828  
Db 5015 TAATGAAGAGAGAAACTACCGAGGAGTCCATAGGATGGCAAGATCTGGTATCGGTC 5074  
QY 4829 ----- 4828  
Db 5075 TGGATTCCGACTCGTCAACATCAATACAACCTATTAAATTTCCCTCGTCAAAAATAAG 5134  
QY 4829 ----- 4828  
Db 5135 GTTATCAAGTGAGAAATCAACATAGTGACGACTGAATCCGGTGAGATGGGCAAAAGCTT 5194  
QY 4829 ----- 4828  
Db 5195 ATGCATTTCTTCAGACTTGTTCACAGCCAGCCATTACGCTCGTCAAAAATCACT 5254  
QY 4829 ----- 4828  
Db 5255 CGCATCAACCAACCGTTATTTCATTCGTGATTCGCTGAGAGAGCAAAATACGGATC 5314  
QY 4829 ----- 4828  
Db 5315 GCTGTTAAAGGACAATTACAAACAGAAATCGAATGCAACCGGGCGAGGAACACTGCCAG 5374  
QY 4829 ----- 4828  
Db 5375 CGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTT 5434  
QY 4829 ----- 4828  
Db 5435 CCCGGGATCCAGTGGTGAGTAACCATGATCATCAGGAGTACGATAAATGCTTGAT 5494  
QY 4829 ----- 4828  
Db 5495 GGTCCGAAGAGCATAAATCCGTCAGCAGCTTTAGTCTGACCATCTCATCTGTAACATC 5554  
QY 4829 ----- 4828  
Db 5555 ATTGGCAACGTACTCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCA 5614  
QY 4829 ----- 4828  
Db 5615 CAATCGATAGATTGTCGACCTGATTGCCGACATATTCGGAGCCATTTATAGCCATA 5674  
QY 4829 ----- 4828  
Db 5675 TAAATCAGCATCCATGTTGGAATTTAATTCGGGCTCGAGCAAGCGTTTCCCGTTGAAT 5734  
QY 4829 ----- 4828  
Db 5735 ATGGCTATAACACCCCTGTTACTGTTTATGTTTATGTAAGCAGACAGTTTATTGTTATGA 5794  
QY 4829 ----- 4828  
Db 5795 TGAATATTTTATCTTGCAATGTAACATCAGAGATTTTGACACAACACGTGGCTTTC 5854  
QY 4829 ----- CTTGAGTTCGGATCACCGGCCACAGGTCGCTGCTGCGCCCTATAT 4879  
Db 5855 CCCCCCCCCCTGACAGTTCGCAATCACCGGCCACAGGTCGCTGCTGCGCCCTATAT 5914  
QY 4880 CGCGACATCAACGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGT 4939  
Db 5915 CGCCGACATCAACGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGT 5974  
QY 4940 CGGCTGGGATGTTGCGAGCCCGCTGCGCGGGGACTGTTGGCGCATCTCCCTGCA 4999  
Db 5975 CGGCTGGGATGTTGCGAGCCCGCTGCGCGGGGACTGTTGGCGCATCTCCCTGCA 6034  
QY 5000 TGCACCATCTCTCGCGCGGGGCTGCTCAACGGGCTCAACCTACTACTGGGCTGCTTCC 5059  
Db 6035 TGCACCATCTCTCGCGCGGGGCTGCTCAACGGGCTCAACCTACTACTGGGCTGCTTCC 6094  
QY 5060 AATGACGAGTCCGATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAAATGGT 5119

Db 6095 AATGACGAGTCCGATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAAATGGT 6154  
QY 5120 GATACCCGATCTCTAGTGTCTTGAAGTCTTCCATCAGATTATGCCCAACTAAGCAAC 5179  
Db 6155 GATACCCGATCTCTAGTGTCTTGAAGTCTTCCATCAGATTATGCCCAACTAAGCAAC 6214  
QY 5180 CGGAGGAGGAGATTTCATGTTAAATTTCTCTGACTTTTGGTTCATCAGTAGACTGCAACTG 5239  
Db 6215 CGGAGGAGGAGATTTCATGTTAAATTTCTCTGACTTTTGGTTCATCAGTAGACTGCAACTG 6274  
QY 5240 TGAGACTATCTCGGTATGACAGCAGAAATGCTCTCTTGGAGACAGTAATGAAGTCCC 5299  
Db 6275 TGAGACTATCTCGGTATGACAGCAGAAATGCTCTCTTGGAGACAGTAATGAAGTCCC 6334  
QY 5300 ACCAATAAGAAATCTTGTATCAGGAACAAATCTTCTTGTTCGAACTTTTTCGGTGCC 5359  
Db 6335 ACCAATAAGAAATCTTGTATCAGGAACAAATCTTCTTGTTCGAACTTTTTCGGTGCC 6394  
QY 5360 TTGAACATAAAATGTAGAGTGATGTCGGGTAGGAATGGAGCGGCAATGCTTACC 5419  
Db 6395 TTGAACATAAAATGTAGAGTGATGTCGGGTAGGAATGGAGCGGCAATGCTTACC 6454  
QY 5420 TTCTGGACCTTCAAGAGTATGAGGTTTGTAGACTACTGATGCCAACTTCAGTGCAAC 5479  
Db 6455 TTCTGGACCTTCAAGAGTATGAGGTTTGTAGACTACTGATGCCAACTTCAGTGCAAC 6514  
QY 5480 GTTGCTATTTCTGTTCAACCAATCCGAATCCAGAGAAATCAAGTTGTTGTCATATT 5539  
Db 6515 GTTGCTATTTCTGTTCAACCAATCCGAATCCAGAGAAATCAAGTTGTTGTCATATT 6574  
QY 5540 GATCCAGCGTCCGCTCTTGAACACTGACAAATAGTGTCTGCTGTTTGGAGTCACTTT 5599  
Db 6575 GATCCAGCGTCCGCTCTTGAACACTGACAAATAGTGTCTGCTGTTTGGAGTCACTTT 6634  
QY 5600 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTGACGAGCAAGGGATATAATACC 5659  
Db 6635 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTGACGAGCAAGGGATATAATACC 6694  
QY 5660 CAATCTAAACTCTTTTAAACGTTAAAGGCAAGTATGTCGCTGTTTAAACCC 5719  
Db 6695 CAATCTAAACTCTTTTAAACGTTAAAGGCAAGTATGTCGCTGTTTAAACCC 6754  
QY 5720 AAATCAGCTCGTAGTCTGATCTCATCAACTTGAGGGGCACTATCTTGTGTTAGAGAAAT 5779  
Db 6755 AAATCAGCTCGTAGTCTGATCTCATCAACTTGAGGGGCACTATCTTGTGTTAGAGAAAT 6814  
QY 5780 TTGCGGAGATCGGATATCGAGAAAGTACGCTGATTTTAAACGTGAATTTATCTCAA 5839  
Db 6815 TTGCGGAGATCGGATATCGAGAAAGTACGCTGATTTTAAACGTGAATTTATCTCAA 6874  
QY 5840 GATCTCTGCTCGCGCTTTCGGTATGACGGTGAACAACTCTGACACATCGAGTCCCG 5899  
Db 6875 GATCTCTGCTCGCGCTTTCGGTATGACGGTGAACAACTCTGACACATCGAGTCCCG 6934  
QY 5900 GAGACGGTCAAGCTTGTCTGTAAGCGGATCCCGGAGCAGACAAGCCGTCAGGGCGG 5959  
Db 6935 GAGACGGTCAAGCTTGTCTGTAAGCGGATCCCGGAGCAGACAAGCCGTCAGGGCGG 6994  
QY 5960 TCAGCGGTGTTGGCGGTTCGGGCGCAGCATGACCCAGTCACTGAGGATAGGGGA 6019  
Db 6995 TCAGCGGTGTTGGCGGTTCGGGCGCAGCATGACCCAGTCACTGAGGATAGGGGA 7054  
QY 6020 GTGTACTCTGCTTAACTATCGGCTCAGAGCAGATTTGTAAGAGTGCACCATATGC 6079  
Db 7055 GTGTACTCTGCTTAACTATCGGCTCAGAGCAGATTTGTAAGAGTGCACCATATGC 7114  
QY 6080 GTGTAAATACCGCAGATGCGTAAAGGAAATACCGCATCAGGCTCTTCCGCTT 6139  
Db 7115 GTGTAAATACCGCAGATGCGTAAAGGAAATACCGCATCAGGCTCTTCCGCTT 7174  
QY 6140 CTTCTCTCACTGACTCGCTGCGCTCGGTGCTGCGCTGCGGCGAGCGGTATCAGTCACT 6199



Db 7175 CCTCGCTCACTGACTCGCTCGGCTCGGTCGTTCCGCTCGGCGAGCGGTATCAGCTCACT 7234  
Qy 6200 CAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATAACGACGAGGAAGAACATGTCAG 6259  
Db 7235 CAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAACATGTCAG 7294  
Qy 6260 CAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAACATGTCAG 6319  
Db 7295 CAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAACATGTCAG 7354  
Qy 6320 GCTCGCGCGCTCGCTGACGAGATCACAAATATCGAGCTCAAGTCAGAGGAGTGGGAAACC 6379  
Db 7355 GCTCGCGCGCTCGCTGACGAGATCACAAATATCGAGCTCAAGTCAGAGGAGTGGGAAACC 7414  
Qy 6380 CGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGTCCCTCGTGCGCTCTCCTG 6439  
Db 7415 CGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGTCCCTCGTGCGCTCTCCTG 7474  
Qy 6440 TTCCGACCGCTCGCTGACGAGTACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCG 6499  
Db 7475 TTCCGACCGCTCGCTGACGAGTACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCG 7534  
Qy 6500 TTCTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGGTTCGGTCCCAAGCTGG 6559  
Db 7535 TTCTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGGTTCGGTCCCAAGCTGG 7594  
Qy 6560 GCTGTGTGCAAGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGTTAACTATCGTC 6619  
Db 7595 GCTGTGTGCAAGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGTTAACTATCGTC 7654  
Qy 6620 TTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGCGCACTGGTAACAGGA 6679  
Db 7655 TTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGCGCACTGGTAACAGGA 7714  
Qy 6680 TTAGCAGAGCGAGGTATGAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCTAACTACG 6739  
Db 7715 TTAGCAGAGCGAGGTATGAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCTAACTACG 7774  
Qy 6740 GCTACACTAGAGACAGTATTTGGTATCTCGCTCTGCTGAGCGAGTACCTTCGAA 6799  
Db 7775 GCTACACTAGAGACAGTATTTGGTATCTCGCTCTGCTGAGCGAGTACCTTCGAA 7834  
Qy 6800 AAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGGTAGCGGTGGTTTTTTG 6859  
Db 7835 AAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGGTAGCGGTGGTTTTTTG 7894  
Qy 6860 TTTCAGACGAGATATACGGCGAGAAAGAAAGATCTCAAGAGATCTCTTGATCTTTT 6919  
Db 7895 TTTCAGACGAGATATACGGCGAGAAAGAAAGATCTCAAGAGATCTCTTGATCTTTT 7954  
Qy 6920 CTACGGGTCTGACGCTCAGTGGAAACAAACTCAGCTTAAGGATTTTGGTCAATGAGAT 6979  
Db 7955 CTACGGGTCTGACGCTCAGTGGAAACAAACTCAGCTTAAGGATTTTGGTCAATGAGAT 8014  
Qy 6980 TATCAAAAGGATCTTCACTTAGATCTCTTTTAAATTAAGTATTAAGTCAATCT 7039  
Db 8015 TATCAAAAGGATCTTCACTTAGATCTCTTTTAAATTAAGTATTAAGTCAATCT 8074  
Qy 7040 AAAGTATATAGTAACTTGGTCTGACAGTACCAATGCTTAATCAGTGAGGACCTTA 7099  
Db 8075 AAAGTATATAGTAACTTGGTCTGACAGTACCAATGCTTAATCAGTGAGGACCTTA 8134  
Qy 7100 TCTCAGCGATCTGTCTATTTGCTTCATCCATAGTTGCTGACTCCCGCTGGTAGATAA 7159  
Db 8135 TCTCAGCGATCTGTCTATTTGCTTCATCCATAGTTGCTGACTCCCGCTGGTAGATAA 8194  
Qy 7160 CTACGATACGGAGGCTTACATCTGGCCCGAGTGTGCAATGATACCGGAGACCCAC 7219  
Db 8195 CTACGATACGGAGGCTTACATCTGGCCCGAGTGTGCAATGATACCGGAGACCCAC 8254  
Qy 7220 GCTCAGCGCTCCAGATTTATCAGCAATACACGAGCGAGCGGAGCGGAGCGAGAA 7279  
Db 8255 GCTCAGCGCTCCAGATTTATCAGCAATACACGAGCGAGCGGAGCGGAGCGGAGCGAGAA 8314

RESULT 5  
US-10-112-221A-11  
; Sequence 11, Application US/10112221A

Qy 7280 GTGCTCTCGCAACTTTTATCCGCTCCATCCAGTCTAATTAATTTGTTCCCGGAGCTAGAG 7339  
Db 8315 GTGCTCTCGCAACTTTTATCCGCTCCATCCAGTCTAATTAATTTGTTCCCGGAGCTAGAG 8374  
Qy 7340 TAAAGTAGTTGCGCAGTTAATAGTTTCGCAACGTTTTCGCAATGCTGTCAGGCAATCGTG 7399  
Db 8375 TAAAGTAGTTGCGCAGTTAATAGTTTCGCAACGTTTTCGCAATGCTGTCAGGCAATCGTG 8434  
Qy 7400 TGTACGCTCGCTGTTTGGTATGGCTTCAATCAGCTCCGGTTCGCTCCGATCGTTG 7459  
Db 8435 TGTACGCTCGCTGTTTGGTATGGCTTCAATCAGCTCCGGTTCGCTCCGATCGTTG 8494  
Qy 7460 TTAATGATCCCGCACTGTTTGTGCAAAAGAGCGGTAGCTTCCGTCCTCGATCGTTG 7519  
Db 8495 TTAATGATCCCGCACTGTTTGTGCAAAAGAGCGGTAGCTTCCGTCCTCGATCGTTG 8554  
Qy 7520 TCAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGAGCACTGCAATATCTC 7579  
Db 8555 TCAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGAGCACTGCAATATCTC 8614  
Qy 7580 TTAATGATCCCGCACTGTTTGTGCAAAAGAGCGGTAGCTTCCGTCCTCGATCGTTG 7639  
Db 8615 TTAATGATCCCGCACTGTTTGTGCAAAAGAGCGGTAGCTTCCGTCCTCGATCGTTG 8674  
Qy 7640 TCTGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGAGCACTGCAATATCTC 7699  
Db 8675 TCTGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGAGCACTGCAATATCTC 8734  
Qy 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATGGAAGAGTCTTCCGCGCGAA 7759  
Db 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATGGAAGAGTCTTCCGCGCGAA 8794  
Qy 7760 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCTGCAACCA 7819  
Db 8795 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCTGCAACCA 8854  
Qy 7820 ACTGATCTTCAGCATCTTTTACCTTCCAGCGTTCCTGGTGAGCAAAACAGGAGCG 7879  
Db 8855 ACTGATCTTCAGCATCTTTTACCTTCCAGCGTTCCTGGTGAGCAAAACAGGAGCG 8914  
Qy 7880 AATAATCCGCAAAAGAGGAAATAGGCGGACACGGAATGTTGAATCTCATACTCTTCC 7939  
Db 8915 AATAATCCGCAAAAGAGGAAATAGGCGGACACGGAATGTTGAATCTCATACTCTTCC 8974  
Qy 7940 TTTTCAATATTTTGAAGCAATTTATCAGGTTTATGCTCTCATGAGCGGATACATTTG 7999  
Db 8975 TTTTCAATATTTTGAAGCAATTTATCAGGTTTATGCTCTCATGAGCGGATACATTTG 9034  
Qy 8000 AATGATTTTGAAGAAATAAACAATAAGGAGTTCCGCGACATTTCCCGGAAAGTCCAC 8059  
Db 9035 AATGATTTTGAAGAAATAAACAATAAGGAGTTCCGCGACATTTCCCGGAAAGTCCAC 9094  
Qy 8060 CTGACGCTTAAGAAACCAATTTATCATGACATTAACCTTAAATAATAGGCGTATCACGA 8119  
Db 9095 CTGACGCTTAAGAAACCAATTTATCATGACATTAACCTTAAATAATAGGCGTATCACGA 9154  
Qy 8120 GGCCTCTTCTCAAGAAATAAACAATAAGGAGTTCCGCGACATTTCCCGGAAAGTCCAC 8179  
Db 9155 GGCCTCTTCTCAAGAAATAAACAATAAGGAGTTCCGCGACATTTCCCGGAAAGTCCAC 9214  
Qy 8180 TCAATGTTGATTTGTAAGTATAGCAGATCGGGAACCACTGAAATAATACAGTTATTT 8239  
Db 9215 TCAATGTTGATTTGTAAGTATAGCAGATCGGGAACCACTGAAATAATACAGTTATTT 9274  
Qy 8240 CG 8241  
Db 9275 CG 9276



Db 1595 GCTTTAATGCGTAGTTTATACAGTTAAATGCTAAGCGAGTCAGGCACCGTGTATGAA 1654  
QY 1861 ATCTAACAAATGCGCTCATGCTCATCTCGGCACCGCTACCCCTGGATGCTGTAGGCATAGG 1920  
Db 1655 ATCTAACAAATGCGCTCATGCTCATCTCGGCACCGCTACCCCTGGATGCTGTAGGCATAGG 1714  
QY 1921 CTTGGTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGGACAGCATCGC 1980  
Db 1715 CTTGGTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGGACAGCATCGC 1774  
QY 1981 CAGTCACATATGCGGTGCTAGCGCTATATGCGGTGATGCAATTTCTATGCGACCGCT 2040  
Db 1775 CAGTCACATATGCGGTGCTAGCGCTATATGCGGTGATGCAATTTCTATGCGACCGCT 1834  
QY 2041 TCTCGAGCAGTGTCCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTCGCTACTTGG 2100  
Db 1835 TCTCGAGCAGTGTCCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTCGCTACTTGG 1894  
QY 2101 AGCCACTATCGACTACGGGATCATGGCGACACACCGCTCTGTGGATCTATCGAATCTA 2160  
Db 1895 AGCCACTATCGACTACGGGATCATGGCGACACACCGCTCTGTGGATCTATCGAATCTA 1954  
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATAGCAACCTTAA 2220  
Db 1955 AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATAGCAACCTTAA 2014  
QY 2221 CAGCAATTCGGGTAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA 2280  
Db 2015 CAGCAATTCGGGTAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA 2074  
QY 2281 TATGTTTCAGTCCCTCAGGAGTTAGTCTTGTGAAGTGAATGAATTTCTGGAAGTGTGCAG 2340  
Db 2075 TATGTTTCAGTCCCTCAGGAGTTAGTCTTGTGAAGTGAATGAATTTCTGGAAGTGTGCAG 2134  
QY 2341 TGTTAACTCCGCTGATTTGAAGGAGATATCGTAGTTGGCAAGTAGTGTGTAGTACGG 2400  
Db 2135 TGTTAACTCCGCTGATTTGAAGGAGATATCGTAGTTGGCAAGTAGTGTGTAGTACGG 2194  
QY 2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGACCAACAAACACAGATCCACGCTGTT 2460  
Db 2195 AGGAGTAATCTCCAACTCTCTGAGAGTAGGACCAACAAACACAGATCCACGCTGTT 2254  
QY 2461 GTACTTGATCAACATAAGAGAAGCATTTCTCGATTTGCGAGATCAAGTTTTCAGGAGCGT 2520  
Db 2255 GTACTTGATCAACATAAGAGAAGCATTTCTCGATTTGCGAGATCAAGTTTTCAGGAGCGT 2314  
QY 2521 ACTGATTCGACATTTCCAAAGCCTGCTCGTAGTTGCAACCGATAGGTTGTAGAGTGTG 2580  
Db 2315 ACTGATTCGACATTTCCAAAGCCTGCTCGTAGTTGCAACCGATAGGTTGTAGAGTGTG 2374  
QY 2581 CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTGACAGCTTGGTTGTGAACAGCAT 2640  
Db 2375 CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTGACAGCTTGGTTGTGAACAGCAT 2434  
QY 2641 CTTCAATTTCTGGCAAGCTCTTGTGCTGTATATCGACAGCCAAAGATCAACCTGGGAAT 2700  
Db 2435 CTTCAATTTCTGGCAAGCTCTTGTGCTGTATATCGACAGCCAAAGATCAACCTGGGAAT 2494  
QY 2701 CAATACCAATGTTTCAAGTGTAGACAGAGGTTCTGAGGCAACGAAATCTGGATCAGCGTATT 2760  
Db 2495 CAATACCAATGTTTCAAGTGTAGACAGAGGTTCTGAGGCAACGAAATCTGGATCAGCGTATT 2554  
QY 2761 TATCAGCAATTAATAGAACTTCAAGAGCCAGCAGGATGTCATTAATCTACACAGGCGTG 2820  
Db 2555 TATCAGCAATTAATAGAACTTCAAGAGCCAGCAGGATGTCATTAATCTACACAGGCGTG 2614  
QY 2821 ATGTGTCAATTTGAACCAATCATCTGGCAGCAGTAAAGAACTGCTTCTGGACCAATA 2880  
Db 2615 ATGTGTCAATTTGAACCAATCATCTGGCAGCAGTAAAGAACTGCTTCTGGACCAATA 2674  
QY 2881 TTTTGTCACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTGGCGC 2940

Db 2675 TTTTGTCACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTCGGGCGC 2734  
QY 2941 CTCTCTGTAGCAGCATACACTTAGCACCAACCTTGTGGGCAACGTAGATGACTTCTGGGG 3000  
Db 2735 CTCTCTGTAGCAGCATACACTTAGCACCAACCTTGTGGGCAACGTAGATGACTTCTGGGG 2794  
QY 3001 TAAGGTTACCATCTCTTCTTAGTGTGGAGATGCAAAAACAAATTTCTTGGCAACGAACTT 3060  
Db 2795 TAAGGTTACCATCTCTTCTTAGTGTGGAGATGCAAAAACAAATTTCTTGGCAACGAACTT 2854  
QY 3061 TGGCAGGAAACCCAGCATCAGGGAAGTGGAGGCAAGATTTGGGTTCCACACAGGAATAT 3120  
Db 2855 TGGCAGGAAACCCAGCATCAGGGAAGTGGAGGCAAGATTTGGGTTCCACACAGGAATAT 2914  
QY 3121 AGAGGCCAACTTTCTCAATAGTCTTGGAAAACGAGAGCAGACTACACAGGGCAAGTCT 3180  
Db 2915 AGAGGCCAACTTTCTCAATAGTCTTGGAAAACGAGAGCAGACTACACAGGGCAAGTCT 2974  
QY 3181 CAACTTGGCAACGCTCTCGTTAGTGTAGTTCATGGAAATTTCTGACGTTATCTATAGAGA 3240  
Db 2975 CAACTTGGCAACGCTCTCGTTAGTGTAGTTCATGGAAATTTCTGACGTTATCTATAGAGA 3034  
QY 3241 GATCAATGGCTCTCTTAAACGTTATCTGGCAATTTGCAATTAAGTTCCTCTGGAAAGGAGCTT 3300  
Db 3035 GATCAATGGCTCTCTTAAACGTTATCTGGCAATTTGCAATTAAGTTCCTCTGGAAAGGAGCTT 3094  
QY 3301 CTAACACAGTGTCTTCAAGGAGCTCCATCAAACTTGGCAGTTAGTCTTAAAGGGCTT 3360  
Db 3095 CTAACACAGTGTCTTCAAGGAGCTCCATCAAACTTGGCAGTTAGTCTTAAAGGGCTT 3154  
QY 3361 TGTCAACATTTTGAAGCAACATTTGTGCACAATTTGGTTTGAATAATTCATTAATCTGTCCG 3420  
Db 3155 TGTCAACATTTTGAAGCAACATTTGTGCACAATTTGGTTTGAATAATTCATTAATCTGTCCG 3214  
QY 3421 TTTTCTGGATAGCAGCAGAGGGCATCTTCAATTTCTGTGAGGAGGCTTAGAAACGT 3480  
Db 3215 TTTTCTGGATAGCAGCAGAGGGCATCTTCAATTTCTGTGAGGAGGCTTAGAAACGT 3274  
QY 3481 CAATTTTGCACAATTTCAATACAGACCTTCAAGAGGAGCTTCTTTAGTTTGGATCTTCTT 3540  
Db 3275 CAATTTTGCACAATTTCAATACAGACCTTCAAGAGGAGCTTCTTTAGTTTGGATCTTCTT 3334  
QY 3541 TAGTTTGTCTCTTGGTGTATCTTGGCTTGGCATCTCTTCTTCTTCTTCTAGTGAACCTTAGGG 3600  
Db 3335 TAGTTTGTCTCTTGGTGTATCTTGGCTTGGCATCTCTTCTTCTTCTTCTAGTGAACCTTAGGG 3394  
QY 3601 ACTTCATATCCAGTTTCTCTCACCTCGTCCAAAGTCAACCGTACTTGGCAGCATCTAA 3660  
Db 3395 ACTTCATATCCAGTTTCTCTCACCTCGTCCAAAGTCAACCGTACTTGGCAGCATCTAA 3454  
QY 3661 CTAATGCAAAATAAAATAAGTCAGCAATTTCCAGGCTATATCTTCTTGGATTTAGCTT 3720  
Db 3455 CTAATGCAAAATAAAATAAGTCAGCAATTTCCAGGCTATATCTTCTTGGATTTAGCTT 3514  
QY 3721 CTGCAAGTTTCAGCTTCTCTCCCTAATTTTAGCGTTCAACAAACCTTCTGTCTCAATA 3780  
Db 3515 CTGCAAGTTTCAGCTTCTCTCCCTAATTTTAGCGTTCAACAAACCTTCTGTCTCAATA 3574  
QY 3781 ACCGTTTGTGTATAAGAACCTTCTGAGGATTCCTTACGATCCCAAGGTGGCTTCCA 3840  
Db 3575 ACCGTTTGTGTATAAGAACCTTCTGAGGATTCCTTACGATCCCAAGGTGGCTTCCA 3634  
QY 3841 TGGCTCTAAGACCTTTTGAATGGCCAAACAGGAAGTGGGTTCCAAAGTGACAGAAACCAA 3900  
Db 3635 TGGCTCTAAGACCTTTTGAATGGCCAAACAGGAAGTGGGTTCCAAAGTGACAGAAACCAA 3694  
QY 3901 CACCTGTTTGTTCACACCAAAATTTTCAGCAGTCTCCATCAAACTCCAAATTCGATACCCA 3960  
Db 3695 CACCTGTTTGTTCACACCAAAATTTTCAGCAGTCTCCATCAAACTCCAAATTCGATACCCA 3754  
QY 3961 GCAACTTTTGTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGACACGAGATT 4020  
Db 3755 GCAACTTTTGTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGACACGAGATT 3814



Db 5975 ||||| CGCGTGGGTATGTTGGCAGGCCCGTGGCGGGGACTGTTGGCGCCATCTCCTTGCA 6034  
Qy 5000 TGCACCAATTCCTTGGCGGGGGTGGCTCAACGGGCTCAACCTACTACTGGGGCTGCTTCCT 5059  
Db 6035 TGCACCAATTCCTTGGCGGGGGTGGCTCAACGGGCTCAACCTACTACTGGGGCTGCTTCCT 6094  
Qy 5060 AATCAGAGTTCGATTAAGGAGAGCGTTCGAGTATCTATGATGGAAGTATGGAATGGT 5119  
Db 6095 AATCAGAGTTCGATTAAGGAGAGCGTTCGAGTATCTATGATGGAAGTATGGAATGGT 6154  
Qy 5120 GATACCCGCAATTCCTCACTGCTTGAAGTTCCTTATCAGATTAATGCCCCAATAAAGCAAC 5179  
Db 6155 GATACCCGCAATTCCTCACTGCTTGAAGTTCCTTATCAGATTAATGCCCCAATAAAGCAAC 6214  
Qy 5180 CGGAGGAGGATTCATGTTAAATTCCTGACTTTTGGTTCATCAGTACTGAACTG 5239  
Db 6215 CGGAGGAGGATTCATGTTAAATTCCTGACTTTTGGTTCATCAGTACTGAACTG 6274  
Qy 5240 TGAGACTATCTCGGTTATGACAGAGAAATGTCCTTCTTGGAGACAGTAAATGAATGCC 5299  
Db 6275 TGAGACTATCTCGGTTATGACAGAGAAATGTCCTTCTTGGAGACAGTAAATGAATGCC 6334  
Qy 5300 ACCAATAAAGAAATCCTTGTATCAGGAACAACTTCTGTTTTCGAACTTTTCGGTGCC 5359  
Db 6335 ACCAATAAAGAAATCCTTGTATCAGGAACAACTTCTGTTTTCGAACTTTTCGGTGCC 6394  
Qy 5360 TTGAACATAAATCTAGAGTGGATATGTCGGGTAGGAATGGAGGGGCAAACTCTTACC 5419  
Db 6395 TTGAACATAAATCTAGAGTGGATATGTCGGGTAGGAATGGAGGGGCAAACTCTTACC 6454  
Qy 5420 TTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGACTGATGCAACTTCAGTGAACAA 5479  
Db 6455 TTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGACTGATGCAACTTCAGTGAACAA 6514  
Qy 5480 GTTGTATTTCTGTTCAACACCAATCCGATCAGAGAAATCAAGTGTGTTGCTACTATT 5539  
Db 6515 GTTGTATTTCTGTTCAACACCAATCCGATCAGAGAAATCAAGTGTGTTGCTACTATT 6574  
Qy 5540 GATCAAGCCAGTGGGCTTTGAACTGCAATAGTGTGCTCGTGTGTTTGAAGTCACTTT 5599  
Db 6575 GATCAAGCCAGTGGGCTTTGAACTGCAATAGTGTGCTCGTGTGTTTGAAGTCACTTT 6634  
Qy 5600 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTACGAGCAAGGCGATTAATACC 5659  
Db 6635 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTACGAGCAAGGCGATTAATACC 6694  
Qy 5660 CAAATCTAAACTCTTTTAAACGTTAAAGCAAGATGATCTGCTGTATTAACCC 5719  
Db 6695 CAAATCTAAACTCTTTTAAACGTTAAAGCAAGATGATCTGCTGTATTAACCC 6754  
Qy 5720 AAATCAGCTGTAGTCTGATCTCTATCAACTTTGAGGGGCACTATCTGTTTGAAGAAAT 5779  
Db 6755 AAATCAGCTGTAGTCTGATCTCTATCAACTTTGAGGGGCACTATCTGTTTGAAGAAAT 6814  
Qy 5780 TTGCGGAGTGGATATCAGAAAGTACGCTGATTTTAAACGTTAAAGTATTAATCTCAA 5839  
Db 6815 TTGCGGAGTGGATATCAGAAAGTACGCTGATTTTAAACGTTAAAGTATTAATCTCAA 6874  
Qy 5840 GATCTCTGCTCGCGGTTTTCGGTGATGACGGTGAAACCTCTGACACATGCACTCCCG 5899  
Db 6875 GATCTCTGCTCGCGGTTTTCGGTGATGACGGTGAAACCTCTGACACATGCACTCCCG 6934  
Qy 5900 GAGACGGTACAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAGCCGTCAGGGCG 5959  
Db 6935 GAGACGGTACAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAGCCGTCAGGGCG 6994  
Qy 5960 TCAGCGGTTGTTGGGGTGTTCGGGGCGCAGCCATGACCCAGTCAAGTACGATAGCGGA 6019  
Db 6995 TCAGCGGTTGTTGGGGTGTTCGGGGCGCAGCCATGACCCAGTCAAGTACGATAGCGGA 7054  
Qy 6020 GTGTATCTGGCTTAACTATGCGGCAATCAGAGCAGATTGTACTGAGGTGACCAATATGC 6079

Db 7055 GTGTATCTGGCTTAACTATGCGGCAATCAGAGCAGATTGTACTGAGAGTGCAACCATATGC 7114  
Qy 6080 GGTGTGAATACCGCACAGATCGTAAGAGAGAAATACCGCATACGGGCTCTTCGGCTT 6139  
Db 7115 GGTGTGAATACCGCACAGATCGTAAGAGAGAAATACCGCATACGGGCTCTTCGGCTT 7174  
Qy 6140 CTTCTGCTACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCACT 6199  
Db 7175 CTTCTGCTACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCACT 7234  
Qy 6200 CAAAGCGGTATACCGTTATCCACAGATCAGGGGATAACCGAGAAAGACATGTGAG 6259  
Db 7235 CAAAGCGGTATACCGTTATCCACAGATCAGGGGATAACCGAGAAAGACATGTGAG 7294  
Qy 6260 CAAAGCGGCAGCAAAAGGCGCAGGAAACCGTAAAGAGCCGCTTGTGGCGTTTTTCCATA 6319  
Db 7295 CAAAGCGGCAGCAAAAGGCGCAGGAAACCGTAAAGAGCCGCTTGTGGCGTTTTTCCATA 7354  
Qy 6320 GGTCTCGCCCTCTGACGAGATCAACAAATTCAGCTCAAGTCAGAGTGGCGGCAAC 6379  
Db 7355 GGTCTCGCCCTCTGACGAGATCAACAAATTCAGCTCAAGTCAGAGTGGCGGCAAC 7414  
Qy 6380 CGACAGACTATAAGATACACAGGCTTCCCTCGGAGCTCCCTCGTGGCTCTCTG 6439  
Db 7415 CGACAGACTATAAGATACACAGGCTTCCCTCGGAGCTCCCTCGTGGCTCTCTG 7474  
Qy 6440 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAAGGTGGCG 6499  
Db 7475 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAAGGTGGCG 7534  
Qy 6500 TTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGTGTAGTTCGTCCTCCACAGCTG 6559  
Db 7535 TTCTCAATGCTCAGCTGTAGTATCTCAGTTCGGTGTAGTTCGTCCTCCACAGCTG 7594  
Qy 6560 GTGTGTGACGAAACCCCTTCAGCCGACCGCTTCGCTTATCCGTAACATATCGTC 6619  
Db 7595 GCTGTGTGACGAAACCCCTTCAGCCGACCGCTTCGCTTATCCGTAACATATCGTC 7654  
Qy 6620 TTGAGTCCAAACCGGTAAAGACAGCTTATCGGCTGACGAGCCATGTTAAACAG 6679  
Db 7655 TTGAGTCCAAACCGGTAAAGACAGCTTATCGGCTGACGAGCCATGTTAAACAG 7714  
Qy 6680 TTAGCAGAGGAGTATGAGCGGTGCTACAGAGTTCGTTGAGTGGTGGCTTAACTAG 6739  
Db 7715 TTAGCAGAGGAGTATGAGCGGTGCTACAGAGTTCGTTGAGTGGTGGCTTAACTAG 7774  
Qy 6740 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTCGGA 6799  
Db 7775 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTCGGA 7834  
Qy 6800 AAAGGTTCGTAGCTCTTGTATCCGCAACAAACACCGCTGCTAGCGTGGTTTTTTG 6859  
Db 7835 AAAGGTTCGTAGCTCTTGTATCCGCAACAAACACCGCTGCTAGCGTGGTTTTTTG 7894  
Qy 6860 TTTCGAAGCAGCAGATTAAGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTTT 6919  
Db 7895 TTTCGAAGCAGCAGATTAAGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTTT 7954  
Qy 6920 CTAAGCGGTCTGACGCTCAGTGAAGCAAACTCAGCTTAAGGATTTTGGTCAATGAT 6979  
Db 7955 CTAAGCGGTCTGACGCTCAGTGAAGCAAACTCAGCTTAAGGATTTTGGTCAATGAT 8014  
Qy 6980 TATCAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTCTTTAAATCAATCT 7039  
Db 8015 TATCAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTCTTTAAATCAATCT 8074  
Qy 7040 AAAGTATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAAATCAGTGGGCACTTA 7099  
Db 8075 AAAGTATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAAATCAGTGGGCACTTA 8134  
Qy 7100 TCTCAGGATCTGCTCTATTTTCTGTTTCAATGCTTCCCTGACCTCCCGCTGCTAGATAA 7159  
Db 8135 TCTCAGGATCTGCTCTATTTTCTGTTTCAATGCTTCCCTGACCTCCCGCTGCTAGATAA 8194

```
QY 7160 CTACGATACGGAGGGGTTACCACTCTGGCCCCAGTGTCTGCAATGATACCGCGAGACCCAC 7219
Db 8195 CTACGATACGGAGGGGTTACCACTCTGGCCCCAGTGTCTGCAATGATACCGCGAGACCCAC 8254

QY 7220 GCTCACCGGCTCAGAGTTTATCAGCAATAAACAGCCAGCGGAGCGGAGCGAGAGAA 7279
Db 8255 GCTCACCGGCTCAGAGTTTATCAGCAATAAACAGCCAGCGGAGCGGAGCGGAGAGAA 8314

QY 7280 GTGTCCTCGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATTTTCGCCGGAAAGCTAGAG 7339
Db 8315 GTGTCCTCGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATTTTCGCCGGAAAGCTAGAG 8374

QY 7340 TAAGTAGTCCCGAGTTAATAGTTTGGCAACGTTGTGTCATGCTGCGAGGATCGTGG 7399
Db 8375 TAAGTAGTCCCGAGTTAATAGTTTGGCAACGTTGTGTCATGCTGCGAGGATCGTGG 8434

QY 7400 TGTCAAGCTCGTCTTTGGTATGGCTTCATTCAAGCTCCGGTTCCTCAACGATCAAGCGAG 7459
Db 8435 TGTCAAGCTCGTCTTTGGTATGGCTTCATTCAAGCTCCGGTTCCTCAACGATCAAGCGAG 8494

QY 7460 TTACATGATCCCGCATGTTGTGCAAAAAGGGTTAGTCTCTTCGGTCTCCGATCGTTG 7519
Db 8495 TTACATGATCCCGCATGTTGTGCAAAAAGGGTTAGTCTCTTCGGTCTCCGATCGTTG 8554

QY 7520 TCAGAAGTAAGTTCGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCTC 7579
Db 8555 TCAGAAGTAAGTTCGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCTC 8614

QY 7580 TTACTGTCATGCCATCCGTAAGTGTCTTTCTGTGACTGTGTGACTCAACCAAGTCAT 7639
Db 8615 TTACTGTCATGCCATCCGTAAGTGTCTTTCTGTGACTGTGTGACTCAACCAAGTCAT 8674

QY 7640 TCTCAGATAGTGTATCGCGGACCGAGTTGCTCTTCGCCGGGTCAACACGGGATAATA 7699
Db 8675 TCTCAGATAGTGTATCGCGGACCGAGTTGCTCTTCGCCGGGTCAACACGGGATAATA 8734

QY 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATATTGGAAACGTTCTTCGGGCGAA 7759
Db 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATATTGGAAACGTTCTTCGGGCGAA 8794

QY 7760 AACTCTCAAGATCTTACCGCTGTGATGATCCAGTTCGATGTAACCACTCGTGCACCCA 7819
Db 8795 AACTCTCAAGATCTTACCGCTGTGATGATCCAGTTCGATGTAACCACTCGTGCACCCA 8854

QY 7820 ACTGATCTTCAGCATCTTTACTTTACACGCGTTTCTGGGTGAGCAAAACAGGAAGGC 7879
Db 8855 ACTGATCTTCAGCATCTTTACTTTACACGCGTTTCTGGGTGAGCAAAACAGGAAGGC 8914

QY 7880 AAAATGCGCAAAAAGGAATAAGGCGGACACGGAATGTTGAATCTCATACTCTTCC 7939
Db 8915 AAAATGCGCAAAAAGGAATAAGGCGGACACGGAATGTTGAATCTCATACTCTTCC 8974

QY 7940 TTTTTCATATTTTGAAGCATTTATCAGGGTTATCTCTCATGAGCGGATACATATTG 7999
Db 8975 TTTTTCATATTTTGAAGCATTTATCAGGGTTATCTCTCATGAGCGGATACATATTG 9034

QY 8000 AATGTATTTAGAAAAATAAACAATAGGGTTTCGGCGCATTTTCCCGGAAAAGTGGCCAC 8059
Db 9035 AATGTATTTAGAAAAATAAACAATAGGGTTTCGGCGCATTTTCCCGGAAAAGTGGCCAC 9094

QY 8060 CTGACGCTTAGAAACCATTTATCATGACATTAACCTATAAAATAGGCGGTATCACGA 8119
Db 9095 CTGACGCTTAGAAACCATTTATCATGACATTAACCTATAAAATAGGCGGTATCACGA 9154

QY 8120 GGCCCTTTTCTGTTCAAGAAATTAATCTCTCATGTTTGAAGCTTATCATCATGAAGTGCAC 8179
Db 9155 GGCCCTTTTCTGTTCAAGAAATTAATCTCTCATGTTTGAAGCTTATCATCATGAAGTGCAC 9214

QY 8180 TCATGTTGGTATTTGTAATAGCGCAGATCGGGAACACTGGAATAATAACAGTTATTAT 8239
Db 9215 TCATGTTGGTATTTGTAATAGCGCAGATCGGGAACACTGGAATAATAACAGTTATTAT 9274

RESULT 6
US-10-104-271-11
; Sequence 11, Application US/10104271
; Publication No. US20030181658A1
; GENERAL INFORMATION:
; APPLICANT: Madison, Edwin
; APPLICANT: Jiunn-Chern Yeh
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING SERINE PROTEASE CVSP14, THE ENCODED
; TITLE OF INVENTION: POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: 24745-1614
; CURRENT APPLICATION NUMBER: US/10104, 271
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/278,166
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
US-10-104-271-11

Query Match 79.6%; Score 6561.6; DB 15; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY 1 AGATCTAACTCCAAAGACGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAG 60
Db 1 AGATCTAACTCCAAAGACGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAG 60

QY 61 GTCCATTTCTACACATAGTCCCAAGCAACGACAGAGGGATACACTAGCAGACAGACCGT 120
Db 61 GTCCATTTCTACACATAGTCCCAAGCAACGACAGAGGGATACACTAGCAGACAGACCGT 120

QY 121 TGCAAAACGAGGACCTCCACCTCTCTTCTCCTCAACACCCACCTTTTGCCTCGAAAAACC 180
Db 121 TGCAAAACGAGGACCTCCACCTCTCTTCTCCTCAACACCCACCTTTTGCCTCGAAAAACC 180

QY 181 AGCCAGTATTGGCTTGATGAGTCCCTCATTCCTCAATTCCTTATTTAGGCTACTA 240
Db 181 AGCCAGTATTGGCTTGATGAGTCCCTCATTCCTCAATTCCTTATTTAGGCTACTA 240

QY 241 ACACCATGACATTTATTAGCCTGTCTATCTGGCCCCCTCTGGCGAGGTTTCATGTTGTTTA 300
Db 241 ACACCATGACATTTATTAGCCTGTCTATCTGGCCCCCTCTGGCGAGGTTTCATGTTGTTTA 300

QY 301 TTTCCGAATGCAACAGCTCCGATTTACACCCGAAACATCACTCCAGATGAGGCTTTCTG 360
Db 301 TTTCCGAATGCAACAGCTCCGATTTACACCCGAAACATCACTCCAGATGAGGCTTTCTG 360

QY 361 AGTGGGGTCAAAATAGTTTTCATGTTCCCAAAATGGCCAAAACCTGCACAGTTTAAACGCT 420
Db 361 AGTGGGGTCAAAATAGTTTTCATGTTCCCAAAATGGCCAAAACCTGCACAGTTTAAACGCT 420

QY 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGTTCGCAATACCGTTTGT 480
Db 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGTTCGCAATACCGTTTGT 480

QY 481 TTGAATGCTTAACGCGCAGTTGGTCAAAAAGAACTTCCAAAAGTTCGCAATACCGTTTGT 540
Db 481 TTGAATGCTTAACGCGCAGTTGGTCAAAAAGAACTTCCAAAAGTTCGCAATACCGTTTGT 540

QY 541 CTTGTTGGTATTGATGACAAATGCTCAAAAATAAATCTCAATTAATGCTTAGCCAGTCT 600
Db 541 CTTGTTGGTATTGATGACAAATGCTCAAAAATAAATCTCAATTAATGCTTAGCCAGTCT 600

QY 601 CTCATGCTCTTGAAACCCCGTGCATCTGCGGAAACGCAATGGGAAACACCCGCT 660
Db 601 CTCATGCTCTTGAAACCCCGTGCATCTGCGGAAACGCAATGGGAAACACCCGCT 660
```









Db	4775	ACTTTTGGTTTGGCCACGGAACGGTCTCGGTTGTGCGGAAGATCGGTGATCTGATTCCTTCA	4834
Qy	4829	-----	4828
Db	4835	ACTCAGCAAAAGTTTCGATTATTTCAACAAGCGCGCTCCGTCAGATCAGCTAATGCT	4894
Qy	4829	-----	4828
Db	4895	CTGCCAGTGTTPACAACCAATTAAACCAATTCTGATTAGAAAACTCATCGAGCATCAAAATG	4954
Qy	4829	-----	4828
Db	4955	AAACTGCAATTTATTTCATATCAGGATTATCAATACCATAATTTTGAANAAGCGTTTCTG	5014
Qy	4829	-----	4828
Db	5015	TAATGAAGGAGAAAACTCACGAGGAGTTTCCATAGGATGCAAGATTCCTGGTATCGGTC	5074
Qy	4829	-----	4828
Db	5075	TGCGATTCCGACTCGTCCAAATCAATACAACCTATTAAATTTCCCTTCGTCAAAAATAAG	5134
Qy	4829	-----	4828
Db	5135	GTATTCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGCAAAAAGCTT	5194
Qy	4829	-----	4828
Db	5195	ATGCATTTCTTTCAGACTTGTTCACAGGCGCCATTAGCTCTGTCATCAAAATCACT	5254
Qy	4829	-----	4828
Db	5255	CGCATCAACCAACGGTTATTTCATTGTTGATTGCGCTGAGCGAGAGCAAAATACGGATC	5314
Qy	4829	-----	4828
Db	5315	GCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACGGCGCAGGAACAACCTGCCAG	5374
Qy	4829	-----	4828
Db	5375	CGCATCAACAATATTTCACCTGATCAGGATATTCTTCTAATACCTTGGATGCTGTTTT	5434
Qy	4829	-----	4828
Db	5435	CCCGGGATCGAGTGTGAGTAACCATCATCATCAGGAGTACGGATATAAATGCTTGAT	5494
Qy	4829	-----	4828
Db	5495	GGTCGGAAGAGGCATAAATTCGGTCAGCCAGTTTGTAGTCTGACCATCTCATCTGTAACATC	5554
Qy	4829	-----	4828
Db	5555	ATTGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATTA	5614
Qy	4829	-----	4828
Db	5615	CAATCGATAGATGTTCGACCTGATGTCGCCAGCATATTCGAGCCCCATTTATACCCATA	5674
Qy	4829	-----	4828
Db	5675	TAAATCAGCATCCATGTTGGAAATTAATCGCGGCTTCGAGCAAGACGTTTCCCGTTGAAT	5734
Qy	4829	-----	4828
Db	5735	ATGGCTCATACACCCCTTGTATTACTGTTTATGTAAGCAGACAGATTTTATTGTTTCATGA	5794
Qy	4829	-----	4828
Db	5795	TGATATATTTTATCTTGTGCATGTAAATGTAATCATCAGAGATTTTGAGACACAACGTGGCTTTC	5854
Qy	4829	-----	4828
Db	5855	CCCCCCCCCTCTGAGGTTCGGATCAACCGGCGCCACAGGTGCGGTTGCTGCGCCTATAT	5914

Qy	4880	CGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTTCGGGCTCATGAGCGCTTGTTT	4939
Db	5915	CGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTTCGGGCTCATGAGCGCTTGTTT	5974
Qy	4940	CGCGGTGGGTATGCTGGCAGGCCCCGTGGCCGGGGAAGTGTGGCGGCATCTCCTTGA	4999
Db	5975	CGCGGTGGGTATGCTGGCAGGCCCCGTGGCCGGGGAAGTGTGGCGGCATCTCCTTGA	6034
Qy	5000	TGCAACATTCCTTTGGCGGGCGGTCTCAACGGCCCTCAACCTACTACTTGGGCTGCTTCT	5059
Db	6035	TGCAACATTCCTTTGGCGGGCGGTCTCAACGGCCCTCAACCTACTACTTGGGCTGCTTCT	6094
Qy	5060	AATGAGAGATCGCATATAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAATGGT	5119
Db	6095	AATGAGAGATCGCATATAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAATGGT	6154
Qy	5120	GATACCCGCAATTCCTTCAAGTCTCTTCAAGTCTCTTATCAGATTTATCCCACTAAAGCAAC	5179
Db	6155	GATACCCGCAATTCCTTCAAGTCTCTTATCAGATTTATCCCACTAAAGCAAC	6214
Qy	5180	CGGAGGAGAGATTTTCATGTTAAATTTCTCTGACCTTTTGGTCTCATCAGTAGACTCGAATG	5239
Db	6215	CGGAGGAGAGATTTTCATGTTAAATTTCTCTGACCTTTTGGTCTCATCAGTAGACTCGAATG	6274
Qy	5240	TGAGACTATCTCGGTTATGACAGCAAGAAATGCTCTTCTTGGAGACAGTAAATGAAGTCCC	5299
Db	6275	TGAGACTATCTCGGTTATGACAGCAAGAAATGCTCTTCTTGGAGACAGTAAATGAAGTCCC	6334
Qy	5300	ACCAATAAGAAATTCCTTGTATCAGGAACAACTCTTCTGTTTCGAACTTTTCGGTCCC	5359
Db	6335	ACCAATAAGAAATTCCTTGTATCAGGAACAACTCTTCTGTTTCGAACTTTTCGGTCCC	6394
Qy	5360	TTGAACTATAAATGTAGAGTGGATATGTCGGTAGGAATGAGCGGCAATGCTTACC	5419
Db	6395	TTGAACTATAAATGTAGAGTGGATATGTCGGTAGGAATGAGCGGCAATGCTTACC	6454
Qy	5420	TTCTGACCTTCAAGAGTATGTAGGTTGTAGATCTATGATGCTCAACTTTCAGTGAACAC	5479
Db	6455	TTCTGACCTTCAAGAGTATGTAGGTTGTAGATCTATGATGCTCAACTTTCAGTGAACAC	6514
Qy	5480	GTGCTATTTCTGTTCAAAACCTTCCGAATCCAGAGAAATCAAAGTTGTTGCTACTATT	5539
Db	6515	GTGCTATTTCTGTTCAAAACCTTCCGAATCCAGAGAAATCAAAGTTGTTGCTACTATT	6574
Qy	5540	GATCCAAAGCAGTCCGCTCTTGAACCTGACATAGTGTCTGCTGCTGTATTAACCCCT	5599
Db	6575	GATCCAAAGCAGTCCGCTCTTGAACCTGACATAGTGTCTGCTGCTGTATTAACCCCT	6634
Qy	5600	TGATGAATAAATCTAGTCTTTGATCTAAATTAATCTTGACGAGCCGAGGATTAATACC	5659
Db	6635	TGATGAATAAATCTAGTCTTTGATCTAAATTAATCTTGACGAGCCGAGGATTAATACC	6694
Qy	5660	CAATCTAAACCTCTTTTAAACGTTTAAAGGACAGTATGCTGCTGTATTAACCCCT	5719
Db	6695	CAATCTAAACCTCTTTTAAACGTTTAAAGGACAGTATGCTGCTGTATTAACCCCT	6754
Qy	5720	AAATCAGTCTGATGTCATCACTTGAAGGCGCACTATCTTGTGTTAGAGAAAT	5779
Db	6755	AAATCAGTCTGATGTCATCACTTGAAGGCGCACTATCTTGTGTTAGAGAAAT	6814
Qy	5780	TTGCGGAGATCGGATATCGAGAAAAGGTACGCTGATTTTAAACGTGAATTTATCTCAA	5839
Db	6815	TTGCGGAGATCGGATATCGAGAAAAGGTACGCTGATTTTAAACGTGAATTTATCTCAA	6874
Qy	5840	GATCTCTGCTCGCGGTTTTCGGTGTATGACGGTGAATAAATCTTGACATGACGCTCCCG	5899
Db	6875	GATCTCTGCTCGCGGTTTTCGGTGTATGACGGTGAATAAATCTTGACATGACGCTCCCG	6934
Qy	5900	GAGACGTCACAGCTTGTCTGTAAGCGGATCGCGGAGCAGCAAGCCGTCAGGCGCG	5959
Db	6935	GAGACGTCACAGCTTGTCTGTAAGCGGATCGCGGAGCAGCAAGCCGTCAGGCGCG	6994





Db	1415	GTCAATTTGTTCTTCTCGTACGAGCTGCTCTGATCAGCCTATCTCGAGCTGATGAA	1474	Db	2495	CAATACCATGTTACGCTTGAGACAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTATT	2554
Qy	1681	TATCTTGGGTAGGGTTGGGAATCAATTCGAGTTTGATGTTTCTTGGTATTCCC	1740	Qy	2761	TATCAGCAATTAATACTAGAACTTCAGAAAGCCAGCAGGCAATGTCATATACTACACAGGGCTG	2820
Db	1475	TATCTTGGGTAGGGTTGGGAATCAATTCGAGTTTGATGTTTCTTGGTATTCCC	1534	Db	2555	TATCAGCAATTAATACTAGAACTTCAGAAAGCCAGCAGGCAATGTCATATACTACACAGGGCTG	2614
Qy	1741	ACTCCTCTTCAGAGTACAGAAAGATTAAGTGAAGAAGTTTCGTTTGTGCAAGCTTATCGATAA	1800	Qy	2821	ATGCTCATTTTGAACCAATCATCTTGGCAGCAGTAACGAATCTGTTTCTCGACCAATA	2880
Db	1535	ACTCCTCTTCAGAGTACAGAAAGATTAAGTGAAGAAGTTTCGTTTGTGCAAGCTTATCGATAA	1594	Db	2615	ATGCTCATTTTGAACCAATCATCTTGGCAGCAGTAACGAATCTGTTTCTCGACCAATA	2674
Qy	1801	GCTTTAATGCGGTAGTTTATCAGAGTTAAATGTCTAAGCAGCAGCAGCAGCGTGTATGAA	1860	Qy	2881	TTTTCTCAGCTTAGAAACAGTTTCTGTTCCGTAGCCATAGCAGCTACTCCCTGGGGCG	2940
Db	1595	GCTTTAATGCGGTAGTTTATCAGAGTTAAATGTCTAAGCAGCAGCAGCAGCGTGTATGAA	1654	Db	2675	TTTTGTCACTTTAGAAACAGTTTCTGTTCCGTAGCCATAGCAGCTACTCCCTGGGGCG	2734
Qy	1861	ATCTAAACAATGCGTCAATCGTCATCCTCGGCAACCGTCACCTGGATGCTGTAGGCATAGG	1920	Qy	2941	CTCCTGCTAGCAGATACACTTTAGCACCAACCTTGTGGGCAACGTFAGATGACTTCTGGGG	3000
Db	1655	ATCTAAACAATGCGTCAATCGTCATCCTCGGCAACCGTCACCTGGATGCTGTAGGCATAGG	1714	Db	2735	CTCCTGCTAGCAGATACACTTTAGCACCAACCTTGTGGGCAACGTFAGATGACTTCTGGGG	2794
Qy	1921	CTTGTTATGCGGTATCTGCGGGGCTCTTGGCGGGATATGCTCAATTCGACAGCATCGC	1980	Qy	3001	TAAGGGTACCATCTCTTCTTAGTGAGATGCAAAACAAATTTCTTGCACACAGCAACT	3060
Db	1715	CTTGTTATGCGGTATCTGCGGGGCTCTTGGCGGGATATGCTCAATTCGACAGCATCGC	1774	Db	2795	TAAGGGTACCATCTCTTCTTAGTGAGATGCAAAACAAATTTCTTGCACACAGCAACT	2854
Qy	1981	CAGTCACATATGCGGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCCGT	2040	Qy	3061	TGGCAGGAACACCCAGCATCAGGGAAGTGGGAAGCAGAAATTCGCGTTTCCACAGGAATAT	3120
Db	1775	CAGTCACATATGCGGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCCGT	1834	Db	2855	TGGCAGGAACACCCAGCATCAGGGAAGTGGGAAGCAGAAATTCGCGTTTCCACAGGAATAT	2914
Qy	2041	TCTCGAGCATGTCCGACCGCTTTGGCGCGCCCGAGTCTGCTCGCTTCTGCTACTTGG	2100	Qy	3121	AGAGGCCAATTTCTCAATAGGTCTTGCACCAACAGAGCAGACTACACAGGGCAAGTCT	3180
Db	1835	TCTCGAGCATGTCCGACCGCTTTGGCGCGCCCGAGTCTGCTCGCTTCTGCTACTTGG	1894	Db	2915	AGAGGCCAATTTCTCAATAGGTCTTGCACCAACAGAGCAGACTACACAGGGCAAGTCT	2974
Qy	2101	AGCCACTATCGACTACGGATCATGCGGACACACCCGCTCGTGATCTATCGATCTA	2160	Qy	3181	CAACTTGCAACGCTCTCCGTAGTTAGCTTCATGGAAATTCCTCGACGTTATCTATAGAGA	3240
Db	1895	AGCCACTATCGACTACGGATCATGCGGACACACCCGCTCGTGATCTATCGATCTA	1954	Db	2975	CAACTTGCAACGCTCTCCGTAGTTAGCTTCATGGAAATTCCTCGACGTTATCTATAGAGA	3034
Qy	2161	AATGTAAGTTAAATCTCTAATAATTAATAAAGTCCAGTTTCTCCATACGAACCTTAA	2220	Qy	3241	GATCAATGGCTCTTAAACGTTATCTGGCAATTCGCAATTAAGTTCTCTGGGAAAGAGGCTT	3300
Db	1955	AATGTAAGTTAAATCTCTAATAATTAATAAAGTCCAGTTTCTCCATACGAACCTTAA	2014	Db	3035	GATCAATGGCTCTTAAACGTTATCTGGCAATTCGCAATTAAGTTCTCTGGGAAAGAGGCTT	3094
Qy	2221	CAGCAATTCGGGTAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA	2280	Qy	3301	CTAACACAGGTGCTTCAAAAGGCTCCATCAAACTTCGCGAGTTAGTTCTTAAAGGGCTT	3360
Db	2015	CAGCAATTCGGGTAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA	2074	Db	3095	CTAACACAGGTGCTTCAAAAGGCTCCATCAAACTTCGCGAGTTAGTTCTTAAAGGGCTT	3154
Qy	2281	TATGTTTCAGTCCCTCAGGAGTTAGCTTCTGTAAGTATGAACTTCTGGAAGTTGCAG	2340	Qy	3361	TGTCACCAATTTGACGAACATTTGTCGAATTTGTTTGAATTCATTAATCTGTTCCG	3420
Db	2075	TATGTTTCAGTCCCTCAGGAGTTAGCTTCTGTAAGTATGAACTTCTGGAAGTTGCAG	2134	Db	3155	TGTCACCAATTTGACGAACATTTGTCGAATTTGTTTGAATTCATTAATCTGTTCCG	3214
Qy	2341	TGTTAACTCCGCTGTATTTGACGGGCATATCCGTACGTTTGGCAAGTTGGTACCGG	2400	Qy	3421	TTTTCTGATAGGACGAGAGGGCATCTTCAATTTCTGTGAGGAGCCCTTAGAAGCT	3480
Db	2135	TGTTAACTCCGCTGTATTTGACGGGCATATCCGTACGTTTGGCAAGTTGGTACCGG	2194	Db	3215	TTTTCTGATAGGACGAGAGGGCATCTTCAATTTCTGTGAGGAGCCCTTAGAAGCT	3274
Qy	2401	AGGAGTATCTCCAACTCTCTGAGAGTAGGCAACCAACACAGATCCAGCGTGT	2460	Qy	3481	CAATTTTGCACAATTAATACGACCTTCAGAAAGGAGCTTCTTATAGGTTTGAATCTTCTT	3540
Db	2195	AGGAGTATCTCCAACTCTCTGAGAGTAGGCAACCAACACAGATCCAGCGTGT	2254	Db	3275	CAATTTTGCACAATTAATACGACCTTCAGAAAGGAGCTTCTTATAGGTTTGAATCTTCTT	3334
Qy	2461	GTACTTGATCAACATAAGAAGACATCTCGATTTGCAAGTCAAGTCTTCAAGGAGCGT	2520	Qy	3541	TAGGTTGTTCTTGGTGTATCTGCTGGCTGGCATCTCCTTTCTTCTAGTACCTTAGGG	3600
Db	2255	GTACTTGATCAACATAAGAAGACATCTCGATTTGCAAGTCAAGTCTTCAAGGAGCGT	2314	Db	3335	TAGGTTGTTCTTGGTGTATCTGCTGGCTGGCATCTCCTTTCTTCTAGTACCTTAGGG	3394
Qy	2521	ACTGATGGAATTTCCAAAGCTCTGCTAGGTTGCAACCGATAGGTTGTAGAGTGTG	2580	Qy	3601	ACTTCATATCCAGGTTTCTCTCCACTCGTCCAAAGTCCACCGTACTTGGCAGCTTAA	3660
Db	2315	ACTGATGGAATTTCCAAAGCTCTGCTAGGTTGCAACCGATAGGTTGTAGAGTGTG	2374	Db	3395	ACTTCATATCCAGGTTTCTCTCCACTCGTCCAAAGTCCACCGTACTTGGCAGCTTAA	3454
Qy	2581	CAATACACTTTCGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTGTGAACAGCAT	2640	Qy	3661	CTAATGCAAAATAAATAAAGTTCAGCAATTCAGGCTATATCTTCTTGAATTAGCTT	3720
Db	2375	CAATACACTTTCGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTGTGAACAGCAT	2434	Db	3455	CTAATGCAAAATAAATAAAGTTCAGCAATTCAGGCTATATCTTCTTGAATTAGCTT	3514
Qy	2641	CTTCAATTCGGAGAGCTCTTGTCTGTCATATCGACGCAACAGAAATCACCTGGGAAT	2700	Qy	3721	CTGCAAGTTTCATACGCTTCTCCCTTAAATTTTAGGGTTCAACAAACTTCGTCGTCAAATA	3780
Db	2435	CTTCAATTCGGAGAGCTCTTGTCTGTCATATCGACGCAACAGAAATCACCTGGGAAT	2494	Db	3515	CTGCAAGTTTCATACGCTTCTCCCTTAAATTTTAGGGTTCAACAAACTTCGTCGTCAAATA	3574
Qy	2701	CAATACCAATGTTACGCTTGACAGAGAGGTCTGAGGCAAGAAATCTGGATCAGCGTATT	2760	Qy	3781	ACCGTTTGGTATAGAAACCTTCTCGAGCAATTCGTCATCGATCCCAAGGTGGCTTCCA	3840
				Db	3575	ACCGTTTGGTATAGAAACCTTCTCGAGCAATTCGTCATCGATCCCAAGGTGGCTTCCA	3634

QY	3841	TGGCTCTAAGACCCCTTTGATTGGGCCAAAAACAGGAAGTGCCTTCAAAGTGACAGAAACCAA	3900
Db	3635	TGGCTCTAAGACCCCTTTGATTGGGCCAAAAACAGGAAGTGCCTTCAAAGTGACAGAAACCAA	3694
QY	3901	CACCTGTTTGTTCAAACACAAAATTTCAAGCAGTCTCCATCACAAATCAAATTCGATACCCA	3960
Db	3695	CACCTGTTTGTTCAAACACAAAATTTCAAGCAGTCTCCATCACAAATCAAATTCGATACCCA	3754
QY	3961	GCAACTTTTTCAGTTTGCCTCAGATGATAGCACCTTTATACCAAAACCGTAGCAGGATTT	4020
Db	3755	GCAACTTTTTCAGTTTGCCTCAGATGATAGCACCTTTATACCAAAACCGTAGCAGGATTT	3814
QY	4021	GGTAGACTCCAGTTTGTGTCCTTATAGCTCCGGAAATAGACTTTTTCGACGAGTACACCA	4080
Db	3815	GGTAGACTCCAGTTTGTGTCCTTATAGCTCCGGAAATAGACTTTTTCGACGAGTACACCA	3874
QY	4081	GGCCCAACAGTAATTAAGACAGTFCAGCCCAAAAGTAGTGAATAGACCAATCGGGCGGT	4140
Db	3875	GGCCCAACAGTAATTAAGACAGTFCAGCCCAAAAGTAGTGAATAGACCAATCGGGCGGT	3934
QY	4141	CAGTAGTCAAGAAGCCCAAAATTTCACTGACAGGGAATTTTTCAGATCTTTCAGAAA	4200
Db	3935	CAGTAGTCAAGAAGCCCAAAATTTCACTGACAGGGAATTTTTCAGATCTTTCAGAAA	3994
QY	4201	GTTTCGTATTCAGTAGTCAATTTGCCGAGCATCAATTAATGGGGATTATACCAAGAACACAG	4260
Db	3995	GTTTCGTATTCAGTAGTCAATTTGCCGAGCATCAATTAATGGGGATTATACCAAGAACACAG	4054
QY	4261	TGSAAGTCAACTCTACCAACTTTTCGGTCTCAGAAAAAGCATAAACAGTTTCTACTCCGC	4320
Db	4055	TGSAAGTCAACTCTACCAACTTTTCGGTCTCAGAAAAAGCATAAACAGTTTCTACTCCGC	4114
QY	4321	CATTAGTAGAACTTTTCAATTCGCCCGAGTGGAGAGAAAAGGCACACGCGATCTAGCAT	4380
Db	4115	CATTAGTAGAACTTTTCAATTCGCCCGAGTGGAGAGAAAAGGCACACGCGATCTAGCAT	4174
QY	4381	TAGCGGGCAAGGATGCAACTTTATCAACAGGGTCTTATAGATAACCTTAGCGCTCGGA	4440
Db	4175	TAGCGGGCAAGGATGCAACTTTATCAACAGGGTCTTATAGATAACCTTAGCGCTCGGA	4234
QY	4441	TCATTCCTTTGGACAACTCTTCTTCGCAAAATCTAGGTCCAAATCACTTCATTGATACCAT	4500
Db	4235	TCATTCCTTTGGACAACTCTTCTTCGCAAAATCTAGGTCCAAATCACTTCATTGATACCAT	4294
QY	4501	TATTGTACAATCTCAGCAAGTTGTCGATCAGCTCCTCAAAATGGTCTCTGTAAACGGATG	4560
Db	4295	TATTGTACAATCTCAGCAAGTTGTCGATCAGCTCCTCAAAATGGTCTCTGTAAACGGATG	4354
QY	4561	ACTCAACTTGACATTAACCTTGAAGCTCAGTCGATTGAGTGAACCTTGAAGTGTGCA	4620
Db	4355	ACTCAACTTGACATTAACCTTGAAGCTCAGTCGATTGAGTGAACCTTGAAGTGTGCA	4414
QY	4621	GCTGGTCAGCAGCATAGGGAACACGGCTTTTCTACCAACTCAAGGAATTTATCAACT	4680
Db	4415	GCTGGTCAGCAGCATAGGGAACACGGCTTTTCTACCAACTCAAGGAATTTATCAACT	4474
QY	4681	CTGCAACACTTTCGGTATCGAGTAGCAAGGGAATGTCACTTTGAAGTCGGACAGTGAG	4740
Db	4475	CTGCAACACTTTCGGTATCGAGTAGCAAGGGAATGTCACTTTGAAGTCGGACAGTGAG	4534
QY	4741	TGTAGTCTTGAGAAATCTGAAGCCGTAATTTTATATATCATGTAGTCAAGTCAATCAGAGA	4800
Db	4535	TGTAGTCTTGAGAAATCTGAAGCCGTAATTTTATATATCATGTAGTCAAGTCAATCAGAGA	4594
QY	4801	TCCTCTACGCGGACGCATCGTGGCCGA-----	4828
Db	4595	TCCTCTACGCGGACGCATCGTGGCCGAACCTTCGAGGGGGGGGGGGGGGGCTGAGGTCGC	4654
QY	4829	-----	4828
Db	4655	CTCGTGAAGAAAGTGTGTGCTGACTCATACAGGCTGAATCGCCCACTCATCCAGCCAGA	4714

[illegible]







482 TGAAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGCATACCGTTTGTG 541  
 481 TGAAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGCATACCGTTTGTG 540  
 542 TTGTTGGTATTGATGACCAATGCTCAAAAAGTCTCAATTAATGCTTAAAGTCAAGTCTC 601  
 541 TTGTTGGTATTGATGACCAATGCTCAAAAAGTCTCAATTAATGCTTAAAGTCAAGTCTC 600  
 602 TCTATCGCTTCTGAACCCCGGTGACCTGTCGGAACGCAATGCGGAAACACCCGCTT 661  
 601 TCTATCGCTTCTGAACCCCGGTGACCTGTCGGAACGCAATGCGGAAACACCCGCTT 660  
 662 TTTGGATGATTAATGATGCTCCACATTTGATGATGCTTCCAGATTTCTGGTGGGAATCTG 721  
 661 TTTGGATGATTAATGATGCTCCACATTTGATGATGCTTCCAGATTTCTGGTGGGAATCTG 720  
 722 CTGATAGCTTAACGTTTCAATGATCAAAATTTAACTGTTCTAAACCCCTACTTTGACAGCAATA 781  
 721 CTGATAGCTTAACGTTTCAATGATCAAAATTTAACTGTTCTAAACCCCTACTTTGACAGCAATA 780  
 782 TATAAACAAGAGAGTGCCTGCTCTTAAACCTTTTATTAATCATATATAGCTTA 841  
 781 TATAAACAAGAGAGTGCCTGCTCTTAAACCTTTTATTAATCATATATAGCTTA 840  
 842 CTTTCAATTAATGCGACTGGTTCCCAATGACAGCTTTTGAATTTTAAAGCACTTTTAAAGCA 901  
 841 CTTTCAATTAATGCGACTGGTTCCCAATGACAGCTTTTGAATTTTAAAGCACTTTTAAAGCA 900  
 902 AACTTGAGAAGATCAAAAACAACTAATTTTCGAAGGATCCAAACGATGAGATTTTCCTT 961  
 901 AACTTGAGAAGATCAAAAACAACTAATTTTCGAAGGATCCAAACGATGAGATTTTCCTT 960  
 962 CAATTTTACTGCACTTTTATTCGAGCATCTCCGATGATGCTCCAGTCAACACTA 1021  
 961 CAATTTTACTGCACTTTTATTCGAGCATCTCCGATGATGCTCCAGTCAACACTA 1020  
 1022 CAACAAGAGATGAACGCAAAATTCGGCTGAAGCTGTCATCGGTACTCAGATTTAG 1081  
 1021 CAACAAGAGATGAACGCAAAATTCGGCTGAAGCTGTCATCGGTACTCAGATTTAG 1080  
 1082 AAGGGGATTTGATGCTGCTTTGCGCATTTTCAACAGACAAATTAACGGGTATTGT 1141  
 1081 AAGGGGATTTGATGCTGCTTTGCGCATTTTCAACAGACAAATTAACGGGTATTGT 1140  
 1142 TTATAAATACTACTATTCAGCATTTGCTGCTAAAGAGAGGGGTATCTCTGAGAAAA 1201  
 1141 TTATAAATACTACTATTCAGCATTTGCTGCTAAAGAGAGGGGTATCTCTGAGAAAA 1200  
 1202 GAGAGGCTGAAGCCGAGG----- 1219  
 1201 GATCGGACCGGGCAGGGGTTTCGGGAAGAGAGGACACCCCAAAAAGCTGACCCCTTTAG 1260  
 1220 -----CCGAGGCCAGGCCAGGCCAG 1242  
 1261 CCTACAAGAGTTTATCCCAATGTTGGCCGAGAGACCTTAGGCGCCAGGAGGTATG 1320  
 1243 GAAGAAACAT----- 1252  
 1321 AAGGGAAGATCTCCAGAAATCCGAGGATTTAAGGAACTCACCCCAATTAACAACCCCG 1380  
 1253 -----GATCATGCCCCCGGAGAGGATAAATTTGTTGGTCCCGCGGTGTC 1299  
 1381 ACATCATATTTAAGGATGAAGAAAAACCGGAGCGGACAGGCTGATGACTCAGAGGTGTA 1440  
 1300 AC-----CGCCAGCAGTGCACGAGAGAGGTTCCTGTTT 1334  
 1441 AGGACAAGTTGAACGCTTTGGCCATCTCGGTGATGAACAGTGGCCAGGAGTGAACCTGC 1500  
 1335 TGATGACAGTGTCCGGGATTTCCCGTGG----- 1362  
 1501 GGGTGACCGGGGCTGGGACGAAGATGGCCACCTCAGAGGAGTCTCTGCACCTACGAGG 1560

1363 ----- 1362  
 1561 GCCGGCAGTGGACATCACACAGTCTGACCGGACCGCAGCAAGTACGGCATGCTGGCCC 1620  
 1363 -----TGTTTCCACCCCATGGCCATCGAGAACACCTCAAGAGAAGAAT 1405  
 1621 GCCTGGCGGTGGAGCGCGCTTCGACTGGGTGTACTAGAGTCCAGGCAATATCCACT 1680  
 1406 GTCCCTTCTACTA-----GTGGCTAGAAATCCCTAGG-----GCGGCCCGGA 1449  
 1681 GCTCGGTGAAGAGCAGAGAACTCGGTGGCGGCAATCGGAGGCTGATTCGGGCCCGGA 1740  
 1450 ATTAATTCGCTTATAGACATGACTGTTCTCAGTTCAAGTTGGGCACCTTACGAGAAGACCG 1509  
 1741 ATTAATTCGCTTATAGACATGACTGTTCTCAGTTCAAGTTGGGCACCTTACGAGAAGACCG 1800  
 1510 GTCTTGTAGATTTCTAATCAAGAGATGTCAGATGCCATTTGCTGAGAGATCGAGCT 1569  
 1801 GTCTTGTAGATTTCTAATCAAGAGATGTCAGATGCCATTTGCTGAGAGATCGAGCT 1860  
 1570 TCAATTTTGAATCTTTTATTTTAACTATATAGTATAGATTTTGTCTCAATTTG 1629  
 1861 TCAATTTTGAATCTTTTATTTTAACTATATAGTATAGATTTTGTCTCAATTTG 1920  
 1630 TTTCTTCTGTAAGCTTGTCTGATCAGCTTCTCGCAGCTGATGATATATCTTGTG 1689  
 1921 TTTCTTCTGTAAGCTTGTCTGATCAGCTTCTCGCAGCTGATGATATATCTTGTG 1980  
 1690 GTAGGGTTTGGGAAATCAATTCGAGTTTGTATTTTCTTGGTATTTTCCACCTCTT 1749  
 1981 GTAGGGTTTGGGAAATCAATTCGAGTTTGTATTTTCTTGGTATTTTCCACCTCTT 2040  
 1750 CAGAGTACAGAAGATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATG 1809  
 2041 CAGAGTACAGAAGATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATG 2100  
 1810 CGGTAGTTTATCACAGTTTAAATTTGCTAACGAGTCAGGCACCGTGTATGAATCTAACAA 1869  
 2101 CGGTAGTTTATCACAGTTTAAATTTGCTAACGAGTCAGGCACCGTGTATGAATCTAACAA 2160  
 1870 TCGGCTCATGTCATCTCGGCACCGTCACTCGATGATCTGATGATGATGATGATGAT 1929  
 2161 TCGGCTCATGTCATCTCGGCACCGTCACTCGATGATCTGATGATGATGATGATGAT 2220  
 1930 GCGGTACTGCGCGGCTTTCGCGGATATCGTCCATTCGACACATCGCCAGTCACTA 1989  
 2221 GCGGTACTGCGCGGCTTTCGCGGATATCGTCCATTCGACACATCGCCAGTCACTA 2280  
 1990 TCGGCTGCTGCTAGCGCTATATGCTTGTATGCAATTTCTATCGCACCCGTTCTCGGAGC 2049  
 2281 TCGGCTGCTGCTAGCGCTATATGCTTGTATGCAATTTCTATCGCACCCGTTCTCGGAGC 2340  
 2050 ACTGTCCGACCGGCTTTCGCGGCTTTCGCGGATATCGTCCATTCGACACATCGCCAGTCACTA 2109  
 2341 ACTGTCCGACCGGCTTTCGCGGCTTTCGCGGATATCGTCCATTCGACACATCGCCAGTCACTA 2400  
 2110 GCACTACGCGATCATGCGGACCAACACCCGCTCTGCTGGATCTATCGAATCTAAATGTAAGT 2169  
 2401 GCACTACGCGATCATGCGGACCAACACCCGCTCTGCTGGATCTATCGAATCTAAATGTAAGT 2460  
 2170 TAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2229  
 2461 TAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2520  
 2230 GGTGAGCTTACAGCTTCAACAGAGCCAGATCCATCACTCTTGGCCAATATCTTCA 2289  
 2521 GGTGAGCTTACAGCTTCAACAGAGCCAGATCCATCACTCTTGGCCAATATCTTCA 2580  
 2290 GTCCCTCAGAGTTACGCTTGTGAAGTGAATGAATCTTGTGAAGTGTGAGTTAACTC 2349  
 2581 GTCCCTCAGAGTTACGCTTGTGAAGTGAATGAATCTTGTGAAGTGTGAGTTAACTC 2640  
 2350 CGCTGTATTGACGGGCATATCGGTAGCTTGGCAAGTGTGTTGGTATCCGGAGGAGTAAT 2409



Db 2641 CCTGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGGTGTACCCGAGAGTAAT 2700  
QY 2410 CTCACAACTCTCTGAGAGTAGGCAACCAACAAACACAGATCCAGCGTGTGTACTTGAT 2469  
Db 2701 CTCACAACTCTCTGAGAGTAGGCAACCAACAAACACAGATCCAGCGTGTGTACTTGAT 2760  
QY 2470 CAAATAGAGAGAGCAATCTCGATTTCCAGGATCAAGTGTTCAGGAGCGTACTGATGG 2529  
Db 2761 CAAATAGAGAGAGCAATCTCGATTTCCAGGATCAAGTGTTCAGGAGCGTACTGATGG 2820  
QY 2530 ACAATTTCAAAGCCCTGCTGAGTGTGCAACCGATAGGTTGTAGAGTGTGCAATACACT 2589  
Db 2821 ACAATTTCAAAGCCCTGCTGAGTGTGCAACCGATAGGTTGTAGAGTGTGCAATACACT 2880  
QY 2590 TCGGTACAAATTTCAACCCCTTGGCACTGACAGCTTGTGTGACAGCACTCTCAATTC 2649  
Db 2881 TCGGTACAAATTTCAACCCCTTGGCACTGACAGCTTGTGTGACAGCACTCTCAATTC 2940  
QY 2650 TGGCAAGCTCCTTGTCTGTCTATATCGACAGCCACAGAAATCACTGGGAATCAATACCAT 2709  
Db 2941 TGGCAAGCTCCTTGTCTGTCTATATCGACAGCCACAGAAATCACTGGGAATCAATACCAT 3000  
QY 2710 GTTACAGTTGAGACAGAGGTTGAGGCAACGAATCTGATTCAGCGTATTTATCAGCAA 2769  
Db 3001 GTTACAGTTGAG- CAGAAAGGTTGAGGCAACGAATCTGATTCAGCGTATTTATCAGCAA 3059  
QY 2770 TAACTAGAACTTCAGAAAGCCAGCAGCATGTCAATCTGATTCAGCGCTCATGTGCAT 2829  
Db 3060 TAACTAGAACTTCAGAAAGCCAGCAGCATGTCAATCTGATTCAGCGCTCATGTGCAT 3119  
QY 2830 TTTGAACCATCATCTTGGCAGCAGTAACGAATCTGTTTCTGGACCAAAATATTTGTAC 2889  
Db 3120 TTTGAACCATCATCTTGGCAGCAGTAACGAATCTGTTTCTGGACCAAAATATTTGTAC 3179  
QY 2890 ACTTAGGACAGTTTCTGTTCCGTAGCCATAGCAGTACTGCTGGCGGCTCTGCTA 2949  
Db 3180 ACTTAGGACAGTTTCTGTTCCGTAGCCATAGCAGTACTGCTGGCGGCTCTGCTA 3239  
QY 2950 GCACATACACTTAGCACCAACTTGTGGCAACAGTAGATCACTTCTGGGTAAGGGTAC 3009  
Db 3240 GCACATACACTTAGCACCAACTTGTGGCAACAGTAGATCACTTCTGGGTAAGGGTAC 3299  
QY 3010 CATCTCTTAGTGGAGATGCAAAAACAAATTTCTTTGCAACCGACAACTTTGGCAGGAA 3069  
Db 3300 CATCTCTTAGTGGAGATGCAAAAACAAATTTCTTTGCAACCGACAACTTTGGCAGGAA 3359  
QY 3070 CACCCAGCATCAGGGAAGTGGAGGCAAGTTGCGTTCCACAGGAATATAGAGGCCAA 3129  
Db 3360 CACCCAGCATCAGGGAAGTGGAGGCAAGTTGCGTTCCACAGGAATATAGAGGCCAA 3419  
QY 3130 CTTTCTCAATAGGTTCTGCAAAACGAGAGCAGACTACACGAGGCAAGTCTCAACTTGCA 3189  
Db 3420 CTTTCTCAATAGGTTCTGCAAAACGAGAGCAGACTACACGAGGCAAGTCTCAACTTGCA 3479  
QY 3190 AGTCTCCGTAGTTAGCTTCATCGAATTTCCGAGCTTACTATAGAGAGCAATGG 3249  
Db 3480 AGTCTCCGTAGTTAGCTTCATCGAATTTCCGAGCTTACTATAGAGAGCAATGG 3539  
QY 3250 CTCTCTTAAAGCTTATCTGCAATTCATAAGTTCTCTGGGAAAGGAGCTTCTAACACAG 3309  
Db 3540 CTCTCTTAAAGCTTATCTGCAATTCATAAGTTCTCTGGGAAAGGAGCTTCTAACACAG 3599  
QY 3310 GTGCTCTTAAAGCGATCTCAAACTTGGCGAGTTAGTTCTAAAGGGCTTTGTACCAT 3369  
Db 3600 GTGCTCTTAAAGCGATCTCAAACTTGGCGAGTTAGTTCTAAAGGGCTTTGTACCAT 3659  
QY 3370 TTTGACGAACATTTGTGCAAAATTTGTGCAATTTCCATATCTGTTCCGTTTCTGGA 3429  
Db 3660 TTTGACGAACATTTGTGCAAAATTTGTGCAATTTCCATATCTGTTCCGTTTCTGGA 3719  
QY 3430 TAGGACGACGAGGCACTTCAATTTCTTGTAGGAGCGCTTAGAAACGTCATTTTGC 3489

Db 3720 TAGGACGACGAGGCACTTCAATTTCTTGTAGGAGCGCTTAGAAACGTCATTTTGC 3779  
QY 3490 ACAATTCAAATPACGACCTTCAGAAAGGACCTCTTTTAGTTTGGATCTCTTTTAGGTTGT 3549  
Db 3780 ACAATTCAAATPACGACCTTCAGAAAGGACCTCTTTTAGTTTGGATCTCTTTTAGGTTGT 3839  
QY 3550 CTTTGTGTATCTCTGCTTGGCATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3609  
Db 3840 CTTTGTGTATCTCTGCTTGGCATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3899  
QY 3610 CCAGGTTTCT 3669  
Db 3900 CCAGGTTTCT 3959  
QY 3670 AATAAATAGTCAGCACATTTCCAGGCTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3729  
Db 3960 AATAAATAGTCAGCACATTTCCAGGCTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4019  
QY 3730 CATCAGCTTCT 3788  
Db 4020 CATCAGCTTCT 4079  
QY 3789 GTATAGAACCTTCTGAGCATTTGCTCTTACGATCCCAAGGTGCTTCCATGGCTCTTA 3848  
Db 4080 GTATAGAACCTTCTGAGCATTTGCTCTTACGATCCCAAGGTGCTTCCATGGCTCTTA 4138  
QY 3849 AGACCTTTGATTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAAACCAACACCTGTT 3908  
Db 4139 AGACCTTTGATTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAAACCAACACCTGTT 4198  
QY 3909 TGTTCACCAAAATTTCAAGCAGTCTCCATCACAATTCGAATCCAGCAACTTT 3968  
Db 4199 TGTTCACCAAAATTTCAAGCAGTCTCCATCACAATTCGAATCCAGCAACTTT 4258  
QY 3969 TGAAGTGTCTCCAGATGACGACCTTTATACCAACCGTGACGAGAGTTGGTAGACT 4028  
Db 4259 TGAAGTGTCTCCAGATGACGACCTTTATACCAACCGTGACGAGAGTTGGTAGACT 4318  
QY 4029 CCAGTTTGTCTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAAC 4088  
Db 4319 CCAGTTTGTCTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAAC 4378  
QY 4089 GAGTAAITAGAGAGTACAGCCCAAGTAGTGAATAGACCATCCGGGCGGTCTAGTAGTC 4148  
Db 4379 GAGTAAITAGAGAGTACAGCCCAAGTAGTGAATAGACCATCCGGGCGGTCTAGTAGTC 4438  
QY 4149 AAAGACGCCAAACAAATTTCACTGACAGGGAACCTTTTTCACATCTTTCAGAAAGTTCGTAT 4208  
Db 4439 AAAGACGCCCAACAAATTTCACTGACAGGGAACCTTTTTCACATCTTTCAGAAAGTTCGTAT 4498  
QY 4209 TCAGTAGTCAATTTCCGAGCATCAATTAATGGGGATTTATACAGAGCAACAGTGGAGTGC 4268  
Db 4499 TCAGTAGTCAATTTCCGAGCATCAATTAATGGGGATTTATACAGAGCAACAGTGGAGTGC 4558  
QY 4269 ACATCTACCAACTTTCCGGTCTCAGAAAAGGATTAACAGTTCTACTCCGCCATTAGTG 4328  
Db 4559 ACATCTACCAACTTTCCGGTCTCAGAAAAGGATTAACAGTTCTACTCCGCCATTAGTG 4618  
QY 4329 AAACTTTTCAAATTCGCCAGTGGAGAAAGGCAAGGCAACGCGATCTAGCATTTAGCGGC 4388  
Db 4619 AAACTTTTCAAATTCGCCAGTGGAGAAAGGCAAGGCAACGCGATCTAGCATTTAGCGGC 4678  
QY 4389 AAGGATGCAACTTTATCAACCGGCTCTATAGATTAACCGCTAGCGCTGGATCATCCCT 4448  
Db 4679 AAGGATGCAACTTTATCAACCGGCTCTATAGATTAACCGCTAGCGCTGGATCATCCCT 4738  
QY 4449 TGGACAACTCTTTCTGCCAAATCTAGGTCCAAATCACTTTCAITGATGATCACTATTGTAC 4508  
Db 4739 TGGACAACTCTTTCTGCCAAATCTAGGTCCAAATCACTTTCAITGATGATCACTATTGTAC 4798  
QY 4509 AACTTGAAGCAAGTTGTGATCAGCTCCTCAAAATTTGCTCTCTGTAAACGATGATCAACT 4568  
Db 4799 AACTTGAAGCAAGTTGTGATCAGCTCCTCAAAATTTGCTCTCTGTAAACGATGATCAACT 4858

QY	4569	TGCAATTAACCTGAAGCTGATCGATTGAGTGAACCTTGATCAGGTTGTCCAGCTGGTCA	4628	QY	4998	-----	4997
DB	4859	TGCACATTAACCTGAAGCTGATCGATTGAGTGAACCTTGATCAGGTTGTCCAGCTGGTCA	4918	DB	5939	CTGTTGAACAGTCTGGAAAGAAATGCATAAGCTTTTGGCCATTCTCCCGGATTCAGTCG	5998
QY	4629	GCAGATAGGAAACACGGCTTTTCCCTACCAAACTCAAGGAATATCAAACTCTGCAACA	4688	QY	4998	-----	4997
DB	4919	GCAGATAGGAAACACGGCTTTTCCCTACCAAACTCAAGGAATATCAAACTCTGCAACA	4978	DB	5999	TCACCTCATSGTGATTTCTCACTTGATAACCTTATTTTGGACGAGGGAATTAATAGGTT	6058
QY	4689	CTTCGGTATGAGTAGCAAGGAAATGTGCATATCTTGAAGTCGGAAGTGAAGTGTCT	4748	QY	4998	-----	4997
DB	4979	CTTCGGTATGAGTAGCAAGGAAATGTGCATATCTTGAAGTCGGAAGTGAAGTGTCT	5038	DB	6059	GTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGA	6118
QY	4749	TGAGAAATCTGAAGCCGTATTTTATATCAGTGAGTCACTCAGGAGATCCTCTAC	4808	QY	4998	-----	4997
DB	5039	TGAGAAATCTGAAGCCGTATTTTATATCAGTGAGTCACTCAGGAGATCCTCTAC	5098	DB	6119	ACTGCTCGGTGAGTTTCTCCTCAITACAGAAACCGCTTTTCAAAAATATGGTATTG	6178
QY	4809	GCCGACGATCGTGGCCGACCTCGAGTGGGATCACCGGCGCCACAGGTGCGGTGCT	4868	QY	4998	-----	4997
DB	5099	GCCGACGATCGTGGCCGACCTCGAGTGGGATCACCGGCGCCACAGGTGCGGTGCT	5158	DB	6179	ATAATCCTGATATGAATAAATGCAATTTCAITGATGCTCGATGAGTTTCTTAATCAG	6238
QY	4869	GCGCCTATATCGCCGATCACCGATGGGAAGATCGGCTCGCCACTTCCGGCTCATG	4928	QY	4998	-----	4997
DB	5159	GCGCCTATATCGCCGATCACCGATGGGAAGATCGGCTCGCCACTTCCGGCTCATG	5218	DB	6239	AATTGGTTAATTGGTTGTAACTGCACTGGCAGAGCAITACGCTGACTTGACGGACCGCGCT	6298
QY	4929	AGCGCTTGTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGCGCC	4988	QY	4998	-----	4997
DB	5219	AGCGCTTGTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGCGCC	5278	DB	6299	TTGTTGAATAAATCGAACTTTTGTGTGAGTTGAAGATCAGATCAGGATCTTCCGACAA	6358
QY	4989	ATCTCCTTG-----	4997	QY	4998	-----	4997
DB	5279	ATCTCCTTGGACCTGACGGGGGGGGGGGAAACCAACCTGTGTCTCAAAATCTGTAT	5338	DB	6359	CCGACGCGTTCCTGGGCAAGCAAAAGTTCAAATAACCAACTGGTCCACTACAACAA	6418
QY	4998	-----	4997	QY	4998	-----	4997
DB	5339	GTACATTTGCACAGATAAATAATATATCATGACAAATPAAACTGTCTGCTTACATAA	5398	DB	6419	AGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGGATTCAGGCGTG	6478
QY	4998	-----	4997	QY	4998	-----	5008
DB	5399	ACAGTAATACAAGGGTGTATGAGCCATATTTCAACGGGAAACGTCTTGTCAAGGCGC	5458	DB	6479	GTATGAGTCAGCAACACCTTCTTTCACGAGGAGACCTCAGCGCCCCCCCCCTGCAGG	6538
QY	4998	-----	4997	QY	5009	CTTTCGGCGGGGTGCTCAACGGCTCAACCTACTACTGGGCTGCTTCTTAATGCAGGA	5068
DB	5459	GATTAAATCCAAACATGATGCTGATTTATATGGGTATTAATGGGTCGCGATATGTCG	5518	DB	6539	TCACACGCGGGGTGCTCAACGGCTCAACCTACTACTGGGCTGCTTCTTAATGCAGGA	6598
QY	4998	-----	4997	QY	5069	GTGCGATAGGAGAGAGCGTCCAGTATCTATGATTCGAAATGGAATGGTATACCCGC	5128
DB	5519	GGCAATCAGGTGCGACAACTCTATCGATTGTATGGAAAGCCGATGCGCCAGAGTTGTTT	5578	DB	6599	GTGCGATAGGAGAGAGCGTCCAGTATCTATGATTCGAAATGGAATGGTATACCCGC	6658
QY	4998	-----	4997	QY	5129	ATTCTTCAGTGTCTTGAAGTCTCCTATCAGATTATGCCCAACTAAAGCAACCGGAGGAG	5188
DB	5579	TGAAACATGGCAAGGTAGCGTTGCGCAATGATGTTACAGATGGTTCAGACTAACT	5638	DB	6659	ATTCTTCAGTGTCTTGAAGTCTCCTATCAGATTATGCCCAACTAAAGCAACCGGAGGAG	6718
QY	4998	-----	4997	QY	5189	AGATTTCATGTTAAATTTCTCTGACTTTTGGTCACTAGTACGACTCGAACTGTGAGACTAT	5248
DB	5639	GGCTACGGAATTTATGCTCTTCCGACCATCAAGCAITTTATCCGTACTCCTCATGATG	5698	DB	6719	AGATTTCATGTTAAATTTCTCTGACTTTTGGTCACTAGTACGACTCGAACTGTGAGACTAT	6778
QY	4998	-----	4997	QY	5249	CTCGGTTATGACAGCAGAAATGTCCTTCTTGAGACAGTAAATGAATGAATGCCCAATAAA	5308
DB	5699	CATGGTTACTCACCCTCGATGCCGCGGAAACAGCAATTCAGGTTATGAGAAATATC	5758	DB	6779	CTCGGTTATGACAGCAGAAATGTCCTTCTTGAGACAGTAAATGAATGAATGCCCAATAAA	6838
QY	4998	-----	4997	QY	5309	GAATTCCTTGTATCAGGACAAACTTCTTGTTCGAACTTTTTCGGTGCCTTGAACCTAT	5368
DB	5759	CTGATTCCAGTGAAATATTTGTATGCGCTGGCAGTGTCTCTCGCGGGTTCATTGCA	5818	DB	6839	GAATTCCTTGTATCAGGACAAACTTCTTGTTCGAACTTTTTCGGTGCCTTGAACCTAT	6898
QY	4998	-----	4997	QY	5369	AAAACTAGAGTGGATATGTCGGGTAGGAATGGAGCGGGCAAAATGCTTACCTTCTGGACC	5428
DB	5819	TTCTCTGTTTGAATGCTCTTTTAAACAGCGATCGCGTATTTCTGCTCGCTCAGCGCAAT	5878	DB	6899	AAAACTAGAGTGGATATGTCGGGTAGGAATGGAGCGGGCAAAATGCTTACCTTCTGGACC	6958
QY	4998	-----	4997	QY	5429	TTCAAGAGGTATGAGGGTTTGTAGTACTGATGCAACTTTCAGTGAACAAGTCTGCTATT	5488
DB	5879	CACGAATGAAACCGTTTGGTTGATGCGAGTATTTTGTATGACGAGCGGTAATGGCTGGC	5938	DB	6959	TTCAAGAGGTATGAGGGTTTGTAGTACTGATGCAACTTTCAGTGAACAAGTCTGCTATT	7018
QY	4998	-----	4997	QY	5489	TCGTTCAAACCAATTCGGAATCCAGAGAAATCAAAGTTGTTGTCTACTATTGATCCAAAGC	5548



QY 7709 ATACGAGAACTTTAAAGTGTCTCATCTTGGAAAAAGTCTTCTCGGGCGGAAAACTCTCAA 7768  
 Db 9239 ATACGAGAACTTTAAAGTGTCTCATCTTGGAAAAAGTCTTCTCGGGCGGAAAACTCTCAA 9298  
 QY 7769 GGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCATCTGTGACCCAACTGATCTT 7828  
 Db 9299 GGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCATCTGTGACCCAACTGATCTT 9358  
 QY 7829 CAGCATCTTTTACTTCCACGAGCTTTCTCGGTGAGCAAAAACAGGAAGCAAAAATGCCG 7888  
 Db 9359 CAGCATCTTTTACTTCCACGAGCTTTCTCGGTGAGCAAAAACAGGAAGCAAAAATGCCG 9418  
 QY 7889 CAAAAAGGGAATAAGGCGGACACGGAATGTTGAATPACTCATCTCTTCTTTTCAAT 7948  
 Db 9419 CAAAAAGGGAATAAGGCGGACACGGAATGTTGAATPACTCATCTCTTCTTTTCAAT 9478  
 QY 7949 ATTATTGAAGATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGATTT 8008  
 Db 9479 ATTATTGAAGATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGATTT 9538  
 QY 8009 AGAAAAATAAATAAGGCGGACACGGAATGTTGAATPACTCATCTCTTCTTTTCAAT 8068  
 Db 9539 AGAAAAATAAATAAGGCGGACACGGAATGTTGAATPACTCATCTCTTCTTTTCAAT 9598  
 QY 8069 AAGAAACCATTTATTCATGATGATTAACCTTATAAATAAGGCGGATACAGGCGGCTTTC 8128  
 Db 9599 AAGAAACCATTTATTCATGATGATTAACCTTATAAATAAGGCGGATACAGGCGGCTTTC 9658  
 QY 8129 GTCTTCAAGATTAATCTCATGTTTGACAGCTTATCATGATAAGCTGACTCATGTTGG 8188  
 Db 9659 GTCTTCAAGATTAATCTCATGTTTGACAGCTTATCATGATAAGCTGACTCATGTTGG 9718  
 QY 8189 TATTGTGAAATAGACGAGATCGGGAACACTGAAAAATAACAGTTATTTCG 8241  
 Db 9719 TATTGTGAAATAGACGAGATCGGGAACACTGAAAAATAACAGTTATTTCG 9771

RESULT 9  
 US-09-883-848A-34  
 ; Sequence 34, Application US/09883848A  
 ; Publication No. US20030022819A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ling, L.  
 ; APPLICANT: Sanicola-Nadel, M.  
 ; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES  
 ; FILE REFERENCE: CIBT-P01-119  
 ; CURRENT APPLICATION NUMBER: US/09/883,848A  
 ; CURRENT FILING DATE: 2001-06-18  
 ; PRIOR APPLICATION NUMBER: 60/211,919  
 ; PRIOR FILING DATE: 2000-06-16  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 34  
 ; LENGTH: 10462  
 ; TYPE: DNA  
 ; ORGANISM: Plasmid pub116  
 US-09-883-848A-34

Query Match 69.3%; Score 5709.8; DB 10; Length 10462;  
 Best Local Similarity 77.4%; P: 0; Mismatches 142; Indels 2221; Gaps 8;  
 Matches 8096; Conservative

QY 2 GATCTAACATCAAGACGAAAGGTTGAATGAAACCTTTTTCGATCCGACATCCACAGG 61  
 Db 1 GATCTAACATCAAGACGAAAGGTTGAATGAAACCTTTTTCGATCCGACATCCACAGG 60  
 QY 62 TCCATTCTCACATAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 121  
 Db 61 TCCATTCTCACATAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 120  
 QY 122 GCAACGCGAGGACCTTCTCTTCTCTCAACACCTTTTGCATCGAAAAACCA 181

Db 121 GCAACGCGAGGACCTTCTCTTCTCTCAACACCTTTTGCATCGAAAAACCA 180  
 QY 182 GCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTTCCAATTCCTTCTATTAGGCTACTAA 241  
 Db 181 GCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTTCCAATTCCTTCTATTAGGCTACTAA 240  
 QY 242 CACCATGATTTATTAGCTGTCTATCTGTGCCCCCTGCGAGGTTTCATGTTGTTTAT 301  
 Db 241 CACCATGATTTATTAGCTGTCTATCTGTGCCCCCTGCGAGGTTTCATGTTGTTTAT 300  
 QY 302 TTCCGAATGCAACAGCTCCGCTTACACCCGAAACATCACTCCAGATGAGGGCTTCTGA 361  
 Db 301 TTCCGAATGCAACAGCTCCGCTTACACCCGAAACATCACTCCAGATGAGGGCTTCTGA 360  
 QY 362 GTGTGGGCTCAAAATAGTTTTCATGTTCCCAATAGGCCCAAACTGACAGTTTAAACGCTG 421  
 Db 361 GTGTGGGCTCAAAATAGTTTTCATGTTCCCAATAGGCCCAAACTGACAGTTTAAACGCTG 420  
 QY 422 TCTTGGAACTTAATATGACAAAGCGTGATCTCATCAAGATGAACCTAAAGTTTGGTTCGT 481  
 Db 421 TCTTGGAACTTAATATGACAAAGCGTGATCTCATCAAGATGAACCTAAAGTTTGGTTCGT 480  
 QY 482 TGAATATGCTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCCATACCGTTTGTG 541  
 Db 481 TGAATATGCTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCCATACCGTTTGTG 540  
 QY 542 TTGTTTGTGATTTGATTGAGCAATGCTCAAAATATATCTCATTAATGCTTAGCGAGTCTC 601  
 Db 541 TTGTTTGTGATTTGATTGAGCAATGCTCAAAATATATCTCATTAATGCTTAGCGAGTCTC 600  
 QY 602 TCTATCGGTTCTGAAACCCCGGTGCACTGTGCCGAAAGCAAAATGGGAAACACACCCGCTT 661  
 Db 601 TCTATCGGTTCTGAAACCCCGGTGCACTGTGCCGAAAGCAAAATGGGAAACACACCCGCTT 660  
 QY 662 TTTGATGATTTATGCAATTTGCTCCACATTTGATGCTTCCAGATTTCTGGTGGGATACTG 721  
 Db 661 TTTGATGATTTATGCAATTTGCTCCACATTTGATGCTTCCAGATTTCTGGTGGGATACTG 720  
 QY 722 CTGATAGCTTAACGTTTCAATGATCAAAATTTAACTGTTCTAAACCCCTACTTTCACAGCAATA 781  
 Db 721 CTGATAGCTTAACGTTTCAATGATCAAAATTTAACTGTTCTAAACCCCTACTTTCACAGCAATA 780  
 QY 782 TATAAACAGAGGAGCTGCCCTGTCTTAAACCTTTTTTTTATCATCATTTATTAGCTTA 841  
 Db 781 TATAAACAGAGGAGCTGCCCTGTCTTAAACCTTTTTTTTATCATCATTTATTAGCTTA 840  
 QY 842 CTTTCATTAATTCGAGCTGTTTCCAAATTCACAGCTTTTGGATTTTAAACGACTTTTAAACGAC 901  
 Db 841 CTTTCATTAATTCGAGCTGTTTCCAAATTCACAGCTTTTGGATTTTAAACGACTTTTAAACGAC 900  
 QY 902 AACTTTGAGAAGATCAAAAAACAACTTAATTTTCGAAGGATCAAAACGATGAGATTTCTCTT 961  
 Db 901 AACTTTGAGAAGATCAAAAAACAACTTAATTTTCGAAGGATCAAAACGATGAGATTTCTCTT 960  
 QY 962 CAATTTTACTGCACTTTTATTCGAGCATCTCCGCAATTCCTCCAGTTCAGCTCCAGTCAACCTA 1021  
 Db 961 CAATTTTACTGCACTTTTATTCGAGCATCTCCGCAATTCCTCCAGTTCAGCTCCAGTCAACCTA 1020  
 QY 1022 CAACAGAAGATGAAACGGCACAATAATTCGGGCTGAAGCTGTCTCATCGGTTTACTCAGATTAG 1081  
 Db 1021 CAACAGAAGATGAAACGGCACAATAATTCGGGCTGAAGCTGTCTCATCGGTTTACTCAGATTAG 1080  
 QY 1082 AAGGGGATTTGATGTTGCTGTTTGGCCATTTTCCAGCAGCAAAATACCGGTTATTGT 1141  
 Db 1081 AAGGGGATTTGATGTTGCTGTTTGGCCATTTTCCAGCAGCAAAATACCGGTTATTGT 1140  
 QY 1142 TTATAATTAATACTACTTATTCGAGCATTTGCTGTAAAGAAAGAGGGGTATCTCTCGAGAAA 1201  
 Db 1141 TTATAATTAATACTACTTATTCGAGCATTTGCTGTAAAGAAAGAGGGGTATCTCTCGAGAAA 1200  
 QY 1202 GAGAGG----- 1207  
 Db 1201 GATGGGACCGGGCAGGGGTTTCGGGAAGAGAGGAGCACCACCAAAAGCTGACCCCTTTAG 1260

QY 1208 ----- 1207  
Db 1261 CCTACAGCAGTTTATCCCAATGTGGCCGAGAGACCCCTAGGCGCCAGCGAAGTATG 1320  
QY 1208 ----- 1207  
Db 1321 AAGGGAAGATCTCCAGAACTCCGAGCGATTAAAGGAACTCACCCCAATTACAAACCCG 1380  
QY 1208 ----- 1207  
Db 1381 ACATCATATTAAAGGATGAAGAAACACCGAGCGGACAGCGTGATGACTCAGAGGTGTA 1440  
QY 1208 ----- 1207  
Db 1441 AGGACAAAGTTGAAACGCTTTGGCCATCTCGGTGATGAACCACTGGCCAGGAGTGAACCTGC 1500  
QY 1208 ----- 1207  
Db 1501 GGGTGACCGAGGCTGGGACGAAGATGCCACCACTCAGAGGAGTCTCTGCACTACGAGG 1560  
QY 1208 ----- 1207  
Db 1561 GCCGCGAGTGGACATCACCACTGCTGACCGGACCGCAGCAAGTACGGCATGTGGCCC 1620  
QY 1208 ----- 1207  
Db 1621 GCCTGGCGGTGGAGCCCGCTTCGACTGGGTGTACTACGAGTCCAAAGSCACATATCCACT 1680  
QY 1208 ----- 1207  
Db 1681 GCTCGGTGAAAGCAGAGAACTCGGTGGCGGCCAAATCGGAGGCGTGCAGAAAACCTCACA 1740  
QY 1208 ----- 1207  
Db 1741 CATGCCACCGTGCCCGCAGCCTGAACTCTTGGGGGAGCCGTGCTCTCTCTCCCCC 1800  
QY 1208 ----- 1207  
Db 1801 CAAAACCCAAAGACACCTCATGATCTCCCGGACCCCTGAGGTACATGCTGCTGGTGG 1860  
QY 1232 CCCAGGCCAGAGAAACATGATCAGTGGCCCCCGG----- 1269  
Db 1861 ACGTGAGCCACGAAGACCTGAGGTCAAGTTCAACTGTGACGTGACGGCGCGGTGAGGTGC 1920  
QY 1270 ----- 1294  
Db 1921 ATAAATGCCAAGACAAAGCCGCGGAGGAGCAGTACCAGACGCTACCGTGTGTGTCAGCG 1980  
QY 1295 GTGTACCGCCCGCAGCA----- 1310  
Db 1981 TCCTCACCGTCTGCAACCCAGCACTGGCTGAATGGCAAGAGTACAAGTGAAGTCTTCCA 2040  
QY 1311 ----- 1310  
Db 2041 ACAAAAGCCCTCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGCGAGCCCCGAG 2100  
QY 1311 ----- 1310  
Db 2101 AACCAAGGTGTACACCCCTGCCCCCATCCCGGATGAGTGCACCAAGAACCAAGTCAAGCC 2160  
QY 1311 ----- 1366  
Db 2161 TGACCTGCTGTCAAGGCTTCTATCCAGCEACATCCCGTGGAGTGGGAGAGCAATG 2220  
QY 1367 TCCACCCCATGGCCATCGA----- 1385  
Db 2221 GCGAGCCGAGAGAACTACAAGCCAGCCCTCCCGTGTGGACTCCGAGCGCTCCTTCT 2280  
QY 1386 ----- 1385  
Db 2281 TCCTCTACAGAACTCACCGTGGACAAAGAGCAGGTGGCAGCAGGGAAACGTTCTTCTCAT 2340

QY 1386 ----- 1423  
Db 2341 GCTCCGTGATGATGAGGCTCTGCACACCACTACACGACAGAGGCTCTCCCTGTCTC 2400  
QY 1424 CFTAGAAATCCCTAGGCGCGCCGAATTAATTCGCTTTPAGACATGTTCTCTCAGTT 1483  
Db 2401 CCGGAAATGAGTGCAGCGCGCCGAATTAATTCGCTTTPAGACATGTTCTCTCAGTT 2460  
QY 1484 CRAAGTTGGCAGCTTACGAGAACCGGCTCTCTAGATTCTAATCAAGAGGATGTGAGAA 1543  
Db 2461 CAAAGTTGGGACCTTACGAGAACCGGCTCTCTAGATTCTAATCAAGAGGATGTGAGAA 2520  
QY 1544 TGCATTTGCTGAGAGATGCGAGCTTCAATTTTGTACTTTTTTATTTGTAACCTATAT 1603  
Db 2521 TGCATTTGCTGAGAGATGCGAGCTTCAATTTTGTACTTTTTTATTTGTAACCTATAT 2580  
QY 1604 AGTATAGGATTTTTTTTGTCAATTTTGTCTCTGTAGAGCTTGTCTCTCATCAGCT 1663  
Db 2581 AGTATAGGATTTTTTTTGTCAATTTTGTCTCTGTAGAGCTTGTCTCTCATCAGCT 2640  
QY 1664 ATCTCCAGCTGATGAATATCTTGTGTAGGGGTTTGGGAAATCATTCGAGTTTGTATGT 1723  
Db 2641 ATCTCCAGCTGATGAATATCTTGTGTAGGGGTTTGGGAAATCATTCGAGTTTGTATGT 2700  
QY 1724 TTTTCTGTGATTTTCCCATCTCTCTTCAGAGTACAGAAATTAAGTGAGAGTTCGTTTG 1783  
Db 2701 TTTTCTGTGATTTTCCCATCTCTCTTCAGAGTACAGAAATTAAGTGAGAGTTCGTTTG 2760  
QY 1784 TGCAGCTTATCGATAAGCTTTAATCCGGTAGTTTATCACAGTTAAATTTGCTAAACAGT 1843  
Db 2761 TGCAGCTTATCGATAAGCTTTAATCCGGTAGTTTATCACAGTTAAATTTGCTAAACAGT 2820  
QY 1844 CAGGCACCGTGTATGAAATCTAAACAATGCGCTCATCGTCACTCTCGGCAACCGTCACTG 1903  
Db 2821 CAGGCACCGTGTATGAAATCTAAACAATGCGCTCATCGTCACTCTCGGCAACCGTCACTG 2880  
QY 1904 GATGCTTAGGCTAGGCTTGTATGCGGCTAGTTCGCGGCTCTTGGGGGATATCGTC 1963  
Db 2881 GATGCTTAGGCTAGGCTTGTATGCGGCTAGTTCGCGGCTCTTGGGGGATATCGTC 2940  
QY 1964 CATTCGACAGACATCGCCAGTCACTATGCGCTGCTAGCGCTATATGCTTGTATGCA 2023  
Db 2941 CATTCGACAGACATCGCCAGTCACTATGCGCTGCTAGCGCTATATGCTTGTATGCA 3000  
QY 2024 TTTCTATGCGCACCGTTCGAGGACATGTCGACCGCTTTGGCGCGCGCCAGTCCCTG 2083  
Db 3001 TTTCTATGCGCACCGTTCGAGGACATGTCGACCGCTTTGGCGCGCGCCAGTCCCTG 3060  
QY 2084 CTGCTTTCGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACCAACACCGCTCTG 2143  
Db 3061 CTGCTTTCGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACCAACACCGCTCTG 3120  
QY 2144 TGGATCTATCGAAATCTAAATGTAAAGTTAAATCTCTAAATTAATTAATTAAGTCCCAAT 2203  
Db 3121 TGGATCTATCGAAATCTAAATGTAAAGTTAAATCTCTAAATTAATTAATTAAGTCCCAAT 3180  
QY 2204 CTCCTACGAACTTAAACGATTCGCTGAGCTATAGACCTTCAACAGCAGCCAGATC 2263  
Db 3181 CTCCTACGAACTTAAACGATTCGCTGAGCTATAGACCTTCAACAGCAGCCAGATC 3240  
QY 2264 CATCACTGCTTGGCCAAATATGTTTCACTCCCTCAGGAGTACGCTTTGTGAAGTATGAA 2323  
Db 3241 CATCACTGCTTGGCCAAATATGTTTCACTCCCTCAGGAGTACGCTTTGTGAAGTATGAA 3300  
QY 2324 CTTCTGGAAGTTGAGTGTAACTCCGCTGTATTGACGGGCATATCCGTACGTTGGCAA 2383  
Db 3301 CTTCTGGAAGTTGAGTGTAACTCCGCTGTATTGACGGGCATATCCGTACGTTGGCAA 3360  
QY 2384 AGTGTGTTGCTGACCGGAGGATTAATCTCCCAACTCTCTGGAGAGTAGGACCAACAA 2443  
Db 3361 AGTGTGTTGCTGACCGGAGGATTAATCTCCCAACTCTCTGGAGAGTAGGACCAACAA 3420  
QY 2444 CACAGATCCAGCGTGTGTACTTGTATCAATGAAGAAAGCATTTCTCGATTTTGCAGAT 2503

DB 3421 CACAGATCCAGCGTGTGTACTTCAATCAACATAAGAAAGACATTCGATTTGCAAGAT 3480  
QY 2504 CAAGTGTTCAGGAGCGTACTGATTTGGACATTTCCAAAGCCCTGCTCGTAGGTTTCAACCGA 2563  
DB 3481 CAAGTGTTCAGGAGCGTACTGATTTGGACATTTCCAAAGCCCTGCTCGTAGGTTTCAACCGA 3540  
QY 2564 TAGGGTTGTAGGTGTGCAATACACTTTCGCGTACATTTTCAACCCCTTGGCAACTGCAAGC 2623  
DB 3541 TAGGGTTGTAGGTGTGCAATACACTTTCGCGTACATTTTCAACCCCTTGGCAACTGCAAGC 3600  
QY 2624 TTGGTTGTGAACAGCATCTTCAATTCCTGGCAAGCTCTGCTGTGTATCGACAGCCAA 2683  
DB 3601 TTGGTTGTGAACAGCATCTTCAATTCCTGGCAAGCTCTGCTGTGTATCGACAGCCAA 3660  
QY 2684 CAGAAATACCTGGGAATCAATACCATGTTAGCTTTGAGACAGAAAGTCTGAGGCAAGCA 2743  
DB 3661 CAGAAATACCTGGGAATCAATACCATGTTAGCTTTGAGACAGAAAGTCTGAGGCAAGCA 3719  
QY 2744 ATCTGGATCAGCGTATTTATCAGCAATACCTAGCACTTCAAGGCCCCAGCAGCATGTC 2803  
DB 3720 ATCTGGATCAGCGTATTTATCAGCAATACCTAGCACTTCAAGGCCCCAGCAGCATGTC 3779  
QY 2804 AATACTACACAGGGGTGATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAAGCACTG 2863  
DB 3780 AATACTACACAGGGGTGATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAAGCACTG 3839  
QY 2864 GTTTCCTGGACCAATATTTGTACACTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGC 2923  
DB 3840 GTTTCCTGGACCAATATTTGTACACTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGC 3899  
QY 2924 AGCTACTGCTGGGGCCCTCTGCTAGCAGCATACCTTAGCAACCACTTTGTGGGCAAC 2983  
DB 3900 AGCTACTGCTGGGGCCCTCTGCTAGCAGCATACCTTAGCAACCACTTTGTGGGCAAC 3959  
QY 2984 GTAGATGATCTTGGGGTAAAGGTACATCTTCTTAGTGGAGATGCAAAACAAATTC 3043  
DB 3960 GTAGATGATCTTGGGGTAAAGGTACATCTTCTTAGTGGAGATGCAAAACAAATTC 4019  
QY 3044 TTGCAACACAGCACTTTGGGAGAAACACCCAGCATCAGGGAAGTGAAGAGCAGAAATTC 3103  
DB 4020 TTGCAACACAGCACTTTGGGAGAAACACCCAGCATCAGGGAAGTGAAGAGCAGAAATTC 4079  
QY 3104 GGTTCACACAGGAATATAGAGGCCAACTTTCTCAATAGGTTTCCAAAAAGAGAGCAGAC 3163  
DB 4080 GGTTCACACAGGAATATAGAGGCCAACTTTCTCAATAGGTTTCCAAAAAGAGAGCAGAC 4139  
QY 3164 TACACAGGGCAAGTCTCAACTTCAGAGTCTCCGTTAGTGGATTCATGGAAATTCCT 3223  
DB 4140 TACACAGGGCAAGTCTCAACTTCAGAGTCTCCGTTAGTGGATTCATGGAAATTCCT 4199  
QY 3224 GACGTTATCTATAGAGAGATCAATGGCTCTCTTAACGTTATCTGGCAATTCGATTAAGTTC 3283  
DB 4200 GACGTTATCTATAGAGAGATCAATGGCTCTCTTAACGTTATCTGGCAATTCGATTAAGTTC 4259  
QY 3284 CTCTGGGAAAGAGGTTCTTAACAGAGTCTCTTAACGCGNCTCCATCAACCTTGGCAGT 3343  
DB 4260 CTCTGGGAAAGAGGTTCTTAACAGAGTCTCTTAACGCGNCTCCATCAACCTTGGCAGT 4319  
QY 3344 TAGTTCTAAAGGGGTTTGTCAACATTTGACGAACATTTGCGCAATTTGGTTTGAATA 3403  
DB 4320 TAGTTCTAAAGGGGTTTGTCAACATTTGACGAACATTTGCGCAATTTGGTTTGAATA 4379  
QY 3404 TTCCATAATCTGTTCCGTTTCTGATAGGAGCAGGAAGGCGATCTTCAATTTCTTGTGA 3463  
DB 4380 TTCCATAATCTGTTCCGTTTCTGATAGGAGCAGGAAGGCGATCTTCAATTTCTTGTGA 4439  
QY 3464 GGAGGCCCTTAGAAGCGTCAATTTTGCACAAATTCACATTCAGACCTTCAGAGGAGCTCTTT 3523  
DB 4440 GGAGGCCCTTAGAAGCGTCAATTTTGCACAAATTCACATTCAGACCTTCAGAGGAGCTCTTT 4499  
QY 3524 AGTTTGTGATTCCTTTAGGTTGTCTTGGGTATCTCTGGGTGATCTCTGGCTTGGCACTCTTCTCT 3583

DB 4500 AGTTTGTGATTCCTTTAGGTTGTCTTGGTGTATCTCTGGCTTGGCATCTCTCTTCTCT 4559  
QY 3584 TCTAGTACCTTTAGGAGCTTCATATCCAGGTTTCTCTCCACCTCGTCCACGTCACACC 3643  
DB 4560 TCTAGTACCTTTAGGAGCTTCATATCCAGGTTTCTCTCCACCTCGTCCACGTCACACC 4619  
QY 3644 GTACTTTGACACATCTAACCTTAATGCAAAATAAATAAGTCAGACATTTCCCAAGGTATATC 3703  
DB 4620 GTACTTTGACACATCTAACCTTAATGCAAAATAAATAAGTCAGACATTTCCCAAGGTATATC 4679  
QY 3704 TTTCTTTGATTTAGTCTTCGAAAGTTCATCAGCTTCTCTCTCTTAATTTTAGCGTTC-AAACA 3762  
DB 4680 TTTCTTTGATTTAGTCTTCGAAAGTTCATCAGCTTCTCTCTCTTAATTTTAGCGTTC-AAACA 4739  
QY 3763 AAACCTTCGTCGTCAAATAACCGTTTGTATAGAACCTTCTGGAGCATTTGCTTTACGAT 3822  
DB 4740 AAACCTTCGTCGTCAAATAACCGTTTGTATAGAACCTTCTGGAGCATTTGCTTTACGAT 4799  
QY 3823 CCCAAGAGTGGCTTCCATGGCTCTTAAGACCTTTGATTGGCCAAAACAGAAAGTGGCTT 3882  
DB 4800 CCCAAGAGT-GCTTCCATGGCTCTTAAGACCTTTGATTGGCCAAAACAGAAAGTGGCTT 4858  
QY 3883 CCAAGTGCAGAAACCAACACCTGTTTCTCAACCAAAATTCAGCAGTCTCCATCAC 3942  
DB 4859 CCAAGTGCAGAAACCAACACCTGTTTCTCAACCAAAATTCAGCAGTCTCCATCAC 4918  
QY 3943 AATCCAATTCGATACCCAGCAACTTTTTCAGTTGCTTCCAGATGTAGCACTTTATACACA 4002  
DB 4919 AATCCAATTCGATACCCAGCAACTTTTTCAGTTGCTTCCAGATGTAGCACTTTATACACA 4978  
QY 4003 AACCGTGCAGAGATTTGTTAGCTCCAGTTTGTCTTATAGCTTCCGGAATAGACT 4062  
DB 4979 AACCGTGCAGAGATTTGTTAGCTCCAGTTTGTCTTATAGCTTCCGGAATAGACT 5038  
QY 4063 TTTTTCGAGAGTACACAGGCCCCAACGAGTAATTTAGAAAGTTCAGCCACCAAGTAGTGA 4122  
DB 5039 TTTTTCGAGAGTACACAGGCCCCAACGAGTAATTTAGAAAGTTCAGCCACCAAGTAGTGA 5098  
QY 4123 ATAGACCATTCGGGGCGGTGAGTTCAAAGACGCCAAACAAATTTCTAGCAGGGAAT 4182  
DB 5099 ATAGACCATTCGGGGCGGTGAGTTCAAAGACGCCAAACAAATTTCTAGCAGGGAAT 5158  
QY 4183 TTTTTCGAGATCTTCAGAAAGTTGCTTATTCAGTAGTCAATTTCCGAGCATCAATTAATGGGA 4242  
DB 5159 TTTTTCGAGATCTTCAGAAAGTTGCTTATTCAGTAGTCAATTTCCGAGCATCAATTAATGGGA 5218  
QY 4243 TTTATACAGAGCAACAGTGGAGTTCACATCTACCAACTTTTTCGGGTCTCAGAAAGCAT 4302  
DB 5219 TTTATACAGAGCAACAGTGGAGTTCACATCTACCAACTTTTTCGGGTCTCAGAAAGCAT 5278  
QY 4303 AAACAGTTCTACTACCGCATTTAGTGAACCTTTTCARATCCCGCAGTGGAGAGAAAG 4362  
DB 5279 AAACAGTTCTACTACCGCATTTAGTGAACCTTTTCARATCCCGCAGTGGAGAGAAAG 5338  
QY 4363 GCACAGCCGATATAGCATTAGCGGCAAGGATGCAACTTTTATCAACAGGGTCTCTATAGA 4422  
DB 5339 GCACAGCCGATATAGCATTAGCGGCAAGGATGCAACTTTTATCAACAGGGTCTCTATAGA 5398  
QY 4423 TAAACCTAGCGCTGGGATCATCTTTTGGACNACTCTTTCTGCCNAATCTAGGTCCAAA 4482  
DB 5399 TAAACCTAGCGCTGGGATCATCTTTTGGACNACTCTTTCTGCCNAATCTAGGTCCAAA 5458  
QY 4483 TCACCTTATGTATACATTTATGTACAACTTCAGCAAGTTCGATCAGCTCTCAAAAT 4542  
DB 5459 TCACCTTATGTATACATTTATGTACAACTTCAGCAAGTTCGATCAGCTCTCAAAAT 5518  
QY 4543 GGTCTCTGTAAACGATGATCACTTTCGACATTTAACTTTGAAGTTCAGTTCATTTAGTGA 4602  
DB 5519 GGTCTCTGTAAACGATGATCACTTTCGACATTTAACTTTGAAGTTCAGTTCATTTAGTGA 5578  
QY 4603 ACTTGATCAGGTTTGTGCAAGTTCAGGAGCATAGGAAACACGGCTTTTCTTACCAAC 4662  
DB 5579 ACTTGATCAGGTTTGTGCAAGTTCAGGAGCATAGGAAACACGGCTTTTCTTACCAAC 5638

QY 4663 TCAAGGAATATCAAACTCTGCAACATTTGGTATGAGGTAGCAAGGAATGTCATAC 4722  
Db 5639 TCAAGGAATATCAAACTCTGCAACATTTGGTATGAGGTAGCAAGGAATGTCATAC 5698  
QY 4723 TTGAAGTCGGACAGTGTAGTCTTGAGAAATCTGAAGCCGTATTTTATATCAGT 4782  
Db 5699 TTGAAGTCGGACAGTGTAGTCTTGAGAAATCTGAAGCCGTATTTTATATCAGT 5758  
QY 4783 GAGTCAGTCATCAGGAGATCCCTTACGCGCGACGCGATCGTGGCGGACCTGCGAGTCGGCA 4842  
Db 5759 GAGTCAGTCATCAGGAGATCCCTTACGCGCGACGCGATCGTGGCGGACCTGCGAGTCGGCA 5818  
QY 4843 TCACCGCGCGCACAGTGTGGTGTGCGGCGCTATATCCGCGACATCACCCATCGGGAAG 4902  
Db 5819 TCACCGCGCGCACAGTGTGGTGTGCGGCGCTATATCCGCGACATCACCCATCGGGAAG 5878  
QY 4903 ATCGGGCTCGGCATCTCGGCTCATGAGCGCTTGTTCGGCGGTGATGTCGAGCGC 4962  
Db 5879 ATCGGGCTCGGCATCTCGGCTCATGAGCGCTTGTTCGGCGGTGATGTCGAGCGC 5938  
QY 4963 CCGTGGCGGGGACTGTGGCGCGCATCTCTTCGATGCAC----- 5004  
Db 5939 CCGTGGCGGGGACTGTGGCGCGCATCTCTTCGATGCAC----- 5004  
QY 5005 ----- 5004  
Db 5999 CCACGTTGTCTCAAAATCTCTGATGTTACATTCGACAGATAAAATATATCATCATG 6058  
QY 5005 ----- 5004  
Db 6059 AACATAAAACTGTCTGTCTTACATAAACAGTAATAACAGGGGTGTATGAGCCATATCA 6118  
QY 5005 ----- 5004  
Db 6119 ACGGAAACGCTTCTGCTCAAGCCCGGATTAATTCBAACATGATGCTGATATATGG 6178  
QY 5005 ----- 5004  
Db 6179 GTATAAATGGGCTCGGATTAATTCGGGCAATCAGGTGCGAATCTATCGATTGTATGG 6238  
QY 5005 ----- 5004  
Db 6239 GAAGCCCGATGCGCAGAGTTGTTCTGMAACATGCGAAAGTAGCGTTGCCAATGATGT 6298  
QY 5005 ----- 5004  
Db 6299 TACAGATGAGATGGTCAGACTAACTGGCTGACGGAATTTATGCTCTTCCGACCATCAA 6358  
QY 5005 ----- 5004  
Db 6359 GCATTTTATCCGTACTCTGTGATGATGTCATGTTACTACCACTGCGATCCCGCGGAAAC 6418  
QY 5005 ----- 5004  
Db 6419 AGCATTCAGGTAATTAGAGAATATCTCTGATTCAGGTGAATATTTGTTGATGCGCTGGC 6478  
QY 5005 ----- 5004  
Db 6479 AGTGTTCCTCGCGGTTGCATTCGATTCCTGTTTGTAAATGTTCTTTTAAACAGCGATCG 6538  
QY 5005 ----- 5004  
Db 6539 CGTATTCGTCTCGCTCAGCGGCAATCACGAATGAATAACGGTTTGGTTGATCGAGTGA 6598  
QY 5005 ----- 5004  
Db 6599 TTTTGATGACGAGGTAATGGCTGGCTGTGTGAACAAGCTGTGAAAGAAATGCATAAGCT 6658  
QY 5005 ----- 5004  
Db 6659 TTTGCCATTCACCGGATTCAGTGTCTACTCATGTGTGATTTCTCATCTTGATAACCTTAT 6718

5005 ----- 5004  
Db 6719 TTTTGACGAGGGAATTAATAGTTGTATTTGATGTTGACGAGTCGGAATCGCAGACCG 6778  
QY 5005 ----- 5004  
Db 6779 ATACCAGGATCTTGCCATCTCTATGGAATGCCTCGGTGAGTTTCTCTCTTACAGAA 6838  
QY 5005 ----- 5004  
Db 6839 ACGGCTTTTCAAAAATATGTTATGATAATCTCTGATATGAATAAATTCGAGTTTCATTT 6898  
QY 5005 ----- 5004  
Db 6899 GATGCTCGATGAGTTTCTTAATCAGAAATGGTTAATTTGGTTGTTAACTGACTGGCAGAGCAT 6958  
QY 5005 ----- 5004  
Db 6959 TACGCTGACTGTGACGGGACGCGCGCTTTGTTGAATAAATTCGAACCTTTTCTGCTGAGTTGAAG 7018  
QY 5005 ----- 5004  
Db 7019 GATCAGATCAGCGCATCTTCCGACAAACGACGACCGTTCCGTGGCAAGCAAAAGTTCAAA 7078  
QY 5005 ----- 5004  
Db 7079 ATCACCAACTGGTCCACCTACAAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGG 7138  
QY 5005 ----- 5004  
Db 7139 CTGGATGATGGGGCGATTGAGGCTGTGATGAGTCAGCAACACCTTCTTCAACGAGGACGA 7198  
QY 5005 -----CATTTCTTGGCGCGGTGCTCAACGGCCTCAACCTTA 5042  
Db 7199 CCTCAGCGCCCGCCCGCTGCGAGGTCCCAACGCGCGCGGTGCTCAACGGCCTCAACCTTA 7258  
QY 5043 CTACTGGGCTGCTTCTTAATGACGAGGTGCTAAGGAGAGCGTGCAGTATCTATGATT 5102  
Db 7259 CTACTGGGCTGCTTCTTAATGACGAGGTGCTAAGGAGAGCGTGCAGTATCTATGATT 7318  
QY 5103 GGAAGTATGGAATGTTGATACCCGCAATCTTCCAGTGTCTTGGTGTCTCTCATCAGATTA 5162  
Db 7319 GGAAGTATGGAATGTTGATACCCGCAATCTTCCAGTGTCTTGGTGTCTCTCATCAGATTA 7378  
QY 5163 TGCCCAACTAAAGCAACCGGAGGAGATTTTCAATGTTAAATTTCTCTGACTTTTGGTCA 5222  
Db 7379 TGCCCAACTAAAGCAACCGGAGGAGATTTTCAATGTTAAATTTCTCTGACTTTTGGTCA 7438  
QY 5223 TCAGTAGACTCGAACTGTGAGACTATCTCGGTTTATGACAGCAAAATGTCTTCTTGGAG 5282  
Db 7439 TCAGTAGACTCGAACTGTGAGACTATCTCGGTTTATGACAGCAAAATGTCTTCTTGGAG 7498  
QY 5283 ACAGTAAATGAAGTCCCAACCAATAAGAAATCTTGTATATGAGAGTATGAGTATCTTGT 5342  
Db 7499 ACAGTAAATGAAGTCCCAACCAATAAGAAATCTTGTATATGAGAGTATGAGTATCTTGT 7558  
QY 5343 CGAATTTTTCGGTGTCTGAACTATATAAATGTAGAGTGTGATATGCGGTAGGAATGGA 5402  
Db 7559 CGAATTTTTCGGTGTCTGAACTATATAAATGTAGAGTGTGATATGCGGTAGGAATGGA 7618  
QY 5403 GCGGGCAAAATGCTTACCTTCTGACCTTCAAGAGGTATGAGGTTTGTATGATCTGATG 5462  
Db 7619 GCGGGCAAAATGCTTACCTTCTGACCTTCAAGAGGTATGAGGTTTGTATGATCTGATG 7678  
QY 5463 CCAATTCAGTGAACAGTTGCTTATTTCTGTTCAACCATTTCCGAATCCAGAGAAATCAAA 5522  
Db 7679 CCAATTCAGTGAACAGTTGCTTATTTCTGTTCAACCATTTCCGAATCCAGAGAAATCAAA 7738  
QY 5523 GTTGTGTTGCTACTATTTGATCCAAAGCCAGTCGCGTCTTGAACCTGAACTAGTGTGCTCG 5582  
Db 7739 GTTGTGTTGCTACTATTTGATCCAAAGCCAGTCGCGTCTTGAACCTGAACTAGTGTGCTCG 7798  
QY 5583 TGTTTTGAGTGCATCTTTTGTATGATAAATCTAGTCTTTGATCTAATAATATCTTGACGAG 5642





[illegible]

RESULT 10

```

US-09-883-848A-32
; Sequence 32, Application US/09883848A
; Publication No. US20030022819A1
; GENERAL INFORMATION:
; APPLICANT: Ling, L.
; APPLICANT: Sanicola-Nadel, M.
; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES
; FILE REFERENCE: CIBT-P01-119
; CURRENT APPLICATION NUMBER: US/09/883,848A
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/211,919
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 10491
; TYPE: DNA
; ORGANISM: Plasmid pUB114
US-09-883-848A-32

Query Match      68.8%; Score 5665.8; DB 10; Length 10491;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 8100; Conservative 0; Mismatches 137; Indels 2252; Gaps 10;

QY      2  GATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCATCCGACATCCACAGG 61
DB      1  GATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCATCCGACATCCACAGG 60

QY      62  TCCATTCTCACATAAGTGCCTAACACGAGGAGGGATACACTAGCAGCAGACCGTT 121
DB      61  TCCATTCTCACATAAGTGCCTAACACGAGGAGGGATACACTAGCAGCAGACCGTT 120

QY      122  GCAACGCGAGGACCTCCACTCTCTCTCTCAACACCCCACTTTTGGCATCGAAAAACA 181
DB      121  GCAACGCGAGGACCTCCACTCTCTCTCTCAACACCCCACTTTTGGCATCGAAAAACA 180

QY      182  GCCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCAATTCCTTTATTAGGCTACTAA 241

```

1243 GAAGAAACAT----- 1252  
QY  
Db 1320 AAGGAAGATCTCAGAAATCCGAGGATTTAAGAACTCACCCCAATTAACAACCCG 1379  
QY  
1253 -----GTATCATGCCCCCGGAGAGGATAAATTTGGCTTCCCGGTGTC 1299  
Db 1380 ACATCATATTTAAGGATGAAGAAAACACCGAGCGGACAGCTGATGACTCAGAGGTGA 1439  
QY  
1300 A----- 1300  
Db 1440 AGACAAGTTGAACGCTTTGGCCATCTCGGTGATGAACCAAGTGGCCAGGAGTGAACATGC 1499  
QY  
1301 ----- 1300  
Db 1500 GGGTGACCGAGGCTGGGACGAAGATGGCCACACTCAGAGGAGTCTCTGCACTACGAGG 1559  
QY  
1301 ----- 1300  
Db 1560 GCCGCGAGTGACATCACCACTGACCGGACCGCAGCAAGTACGGCATGTGGCCC 1619  
QY  
1301 ----- 1300  
Db 1620 GCCTGGCGTGGAGCGCGCTTCGACTCGGTGTACTAGAGTCCAAAGGCACATATCCACT 1679  
QY  
1301 ----- 1300  
Db 1680 GCTCGGTGAAGCAGAGAACTCGGTGGCGGCCAAATCGGAGGCGTGCAGTGGCCAGGG 1739  
QY  
1301 ----- 1300  
Db 1740 ATTGTGTTGTAAGCCTTGCAATGTACAGTCCCGAAGTATCATCTGCTCTCATCTTC 1799  
QY  
1301 ----- 1300  
Db 1800 CCCCCAAGCCAAAGATGTGCTCACCACTTACTCTGACTCTTAAGTCAAGTGTGTGTGG 1859  
QY  
1301 ----- 1300  
Db 1860 TAGACATCAGCAGGATGATCCGAGGTCAGTTTCAGCTGGTTCTAGATGATGTGGAG 1919  
QY  
1301 ----- 1300  
Db 1920 TGCACACAGCTCAGACGCAACACCGGAAGAGCAGTTCCAAAGCACATTTCGCTCAGTCA 1979  
QY  
1301 ----- 1300  
Db 1980 GTGAATCTCCATCATGCACAGGACTGCTCAATGGCAAGAGTTCAATGCAAGGTCA 2039  
QY  
1301 ----- 1300  
Db 2040 ACAGTGCAGCTTTCCTGCCCCCATCGAGAAAACCATCTCCAAAACCAAGGACAGCGA 2099  
QY  
1301 -----CCG 1303  
Db 2100 AGGCTCCACAGGTGACACCATTCACCTCCCAAGGACGATGGCCAGGATAAAGTCA 2159  
QY  
1304 CCCAGCAGTGCCAGGAGAGGTTGCTGTTTGTATGACAGTGTCCGGGGATTCCTCGTGT 1363  
Db 2160 GTCTGACCTGCATGATAACAGACTTCTTCCCTGAAGACATTACTGTGGAGTGGCAGTGA 2219  
QY  
1364 GCTTCAACCCATGGCCATCGAACA----- 1390  
Db 2220 ATGGCAGCCAGCGAGAACTACAAGAACACTCAGCCCATCATGGACACAGATGGCTTT 2279  
QY  
1391 ----- 1390  
Db 2280 ACTTCGTCTACAGCAGCTCAATGTGCAGAGACCACTGGGAGGAGGAGGAAATACTTTCA 2339  
QY  
1391 -----CTCAAGAAAGAAATGTCCTCTTAAT 1418  
Db 2340 CCGTCTGTGTATACATGAGGCGCTGCACAAACCACTACTGAGAGAGCGCTCTCCCACT 2399

1419 AGTGGCGTAGA-----ATTCCCTAGGGCGCGCGCAATTA 1453  
QY  
Db 2400 CTCCTGTAAATGATCCCAAGTGTCTTGGAGCCCTCTGGTCTCCTACAGCGCGCGCAATTA 2459  
QY  
1454 ATTGGCTTTAGACATGACACTGTTCCCTCAGTTCAAGTTGGGCACTTACGAGAACACCGGTCT 1513  
Db 2460 ATTGGCTTTAGACATGACACTGTTCCCTCAGTTCAAGTTGGGCACTTACGAGAACACCGGTCT 2519  
QY  
1514 TGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCCTGAGAGATGAGGTTTCAT 1573  
Db 2520 TGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCCTGAGAGATGAGGTTTCAT 2579  
QY  
1574 TTTTGATACCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTGTTTC 1633  
Db 2580 TTTTGATACCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTGTTTC 2639  
QY  
1634 TTCTCGTACGAGTGTCTCTCTGATCAGCCTATCTCCAGCTGATGAATATCTTTGGTAG 1693  
Db 2640 TTCTCGTACGAGTGTCTCTCTGATCAGCCTATCTCCAGCTGATGAATATCTTTGGTAG 2699  
QY  
1694 GGGTTTGGGAAATCATTCGAGTTTGAATGTTTTCTTGGTATTTCCCACTCCTCTCAGA 1753  
Db 2700 GGGTTTGGGAAATCATTCGAGTTTGAATGTTTTCTTGGTATTTCCCACTCCTCTCAGA 2759  
QY  
1754 GTACAGAAATTAAGTGAGAGTTCGTTTGTCAAGCTTATCGATAAGCTTTAATGCGGT 1813  
Db 2760 GTACAGAAATTAAGTGAGAGTTCGTTTGTCAAGCTTATCGATAAGCTTTAATGCGGT 2819  
QY  
1814 AGTTATCAAGTTAAATTCGTAACGAGTACGGACCGGTGATGAATCTAACAATGCG 1873  
Db 2820 AGTTATCAAGTTAAATTCGTAACGAGTACGGACCGGTGATGAATCTAACAATGCG 2879  
QY  
1874 CTCATCGTATCCTCGGCACCGTCACTCGGTGATGTAGGCATAGGCTTGGTTATGCGG 1933  
Db 2880 CTCATCGTATCCTCGGCACCGTCACTCGGTGATGTAGGCATAGGCTTGGTTATGCGG 2939  
QY  
1934 GTATCGCGGGCTCTTTGGGATATCGTCCATTCGACAGCATCGCAGTCACTATGCG 1993  
Db 2940 GTATCGCGGGCTCTTTGGGATATCGTCCATTCGACAGCATCGCAGTCACTATGCG 2999  
QY  
1994 GTGCTCTAGCCCTATATCGGTTGATGAATTTCTATGGCACCCGTTCTCGAGCACTG 2053  
Db 3000 GTGCTCTAGCCCTATATCGGTTGATGAATTTCTATGGCACCCGTTCTCGAGCACTG 3059  
QY  
2054 TCCGACCGCTTTGGCGCGCGCCAGTCTCTGCTCGTTCTGCTACTTTGGAGCCACTATCGAC 2113  
Db 3060 TCCGACCGCTTTGGCGCGCGCCAGTCTCTGCTCGTTCTGCTACTTTGGAGCCACTATCGAC 3119  
QY  
2114 TAGCGATCATGGCGACACACCGCTCTGTGGATCTATCGAATCTAATGTAAGTTAA 2173  
Db 3120 TAGCGATCATGGCGACACACCGCTCTGTGGATCTATCGAATCTAATGTAAGTTAA 3179  
QY  
2174 ATCTCTAAATTAATAAAGTCCAGATTTCTCCATAGAACTTAAACAGCAATTCGGGTG 2233  
Db 3180 ATCTCTAAATTAATAAAGTCCAGATTTCTCCATAGAACTTAAACAGCAATTCGGGTG 3239  
QY  
2234 AGCATCTAGACCTTCAACAGCAGCAGATCATCATCTGCTTGGCCATATGTTTCAGTCC 2293  
Db 3240 AGCATCTAGACCTTCAACAGCAGCAGATCATCATCTGCTTGGCCATATGTTTCAGTCC 3299  
QY  
2294 CTCAGGAGTTAGCTCTTGTGAAGTATGAATCTCTGGAAGTTGAGTGTAACTCCGCT 2353  
Db 3300 CTCAGGAGTTAGCTCTTGTGAAGTATGAATCTCTGGAAGTTGAGTGTAACTCCGCT 3359  
QY  
2354 GTATTGACGGGATATCCGTACGTTGGCAAGTGTGGTTGTTACCGAGGAGTATCTCC 2413  
Db 3360 GTATTGACGGGATATCCGTACGTTGGCAAGTGTGGTTGTTACCGAGGAGTATCTCC 3419  
QY  
2414 ACNACTCTGAGAGTGGCCACCAACACATCCAGCGTGTGTACTGTATCAAC 2473  
Db 3420 ACNACTCTCTGAGAGTGGCCACCAACACATCCAGCGTGTGTACTGTATCAAC 3479  
QY  
2474 ATAAGAAGAAGCATTTCTCGATTTGCGAGTCAAGTGTTCAGGAGCTACTGATTTGCACAT 2533

3480 ATAAAGAGAGCAATTCCTGATTTGTCAGGATCAAGTTCAGGAGGCTACTGATGGACAT 3539  
2534 TTCCAAAGCCTGCTCGTAGGTTGGAACCGATAGAGGTTGTAGAGTGTGCAATACACTTGGC 2593  
3540 TTCCAAAGCCTGCTCGTAGGTTGGAACCGATAGAGGTTGTAGAGTGTGCAATACACTTGGC 3599  
2594 TACAATTTCAACCCCTTGGCAACTGCACAGCTTGGTGTGAACAGCATCTTCAATTTCTGGC 2653  
3600 TACAATTTCAACCCCTTGGCAACTGCACAGCTTGGTGTGAACAGCATCTTCAATTTCTGGC 3659  
2654 AAGCTCTTGTGCTGTATATGACAGAGCCAAACAGAAATCACTTGGGAATCAATACCATGTTTC 2713  
3660 AAGCTCTTGTGCTGTATATGACAGAGCCAAACAGAAATCACTTGGGAATCAATACCATGTTTC 3719  
2714 AGCTTTGAGCAGAGAGGCTGTAGGCAACGAATCTGGATCAGCGTATTTATCAGCAATAAC 2773  
3720 AGCTTTGAG-CAGAGGCTGTAGGCAACGAATCTGGATCAGCGTATTTATCAGCAATAAC 3778  
2774 TAGAACTTTCAAGAGCCCAAGCAGGATGTCAATATACACAGGGCTGATGTGCTATTTTG 2833  
3779 TAGAACTTTCAAGAGCCCAAGCAGGCAATGTCAATATACACAGGGCTGATGTGCTATTTTG 3838  
2834 AACCATCATCTTGGCAGCGATTAACGAACCTGGTTCCTGGACCAATATTTTGTACACTT 2893  
3839 AACCATCATCTTGGCAGCGATTAACGAACCTGGTTCCTGGACCAATATTTTGTACACTT 3898  
2894 AGGAACAGTTTCTGTTCGTAGGCAATAGCAGCTACTGCTGGGGCCCTCTGCTAGCAC 2953  
3899 AGGAACAGTTTCTGTTCGTAGGCAATAGCAGCTACTGCTGGGGCCCTCTGCTAGCAC 3958  
2954 GATACACTTAGCAACCACTTGTGGGCAACGATAGATGCTTCTGGGGTAAAGGTTACCATC 3013  
3959 GATACACTTAGCAACCACTTGTGGGCAACGATAGATGCTTCTGGGGTAAAGGTTACCATC 4018  
3014 CTTCCTTAGGTGAGATGCAAAACCAATTTCTTTGCAACAGCACTTTGGGAGCAACACC 3073  
4019 CTTCCTTAGGTGAGATGCAAAACCAATTTCTTTGCAACAGCACTTTGGGAGCAACACC 4078  
3074 CAGCATCAGGGAAGTGGAGGAGCAATTCGGTTCACACAGGAATATAGAGGCCAACTTT 3133  
4079 CAGCATCAGGGAAGTGGAGGAGCAATTCGGTTCACACAGGAATATAGAGGCCAACTTT 4138  
3134 CTCAATAGTCTTGAACAGAGAGCAGACTACACAGGCAAGTCTCAACTTGCAGGT 3193  
4139 CTCAATAGTCTTGAACAGAGAGCAGACTACACAGGCAAGTCTCAACTTGCAGGT 4198  
3194 CTCCGTTAGTTGAGCTTCATGGAATTTCTGTAGCTTATCTATAGAGATCAATGGCTCT 3253  
4199 CTCCGTTAGTTGAGCTTCATGGAATTTCTGTAGCTTATCTATAGAGATCAATGGCTCT 4258  
3254 CTTAACGTTATCTGGCAATTCATAGTTCCTCTGGGAAAGAGCTTCTAAACAGGTGT 3313  
4259 CTTAACGTTATCTGGCAATTCATAGTTCCTCTGGGAAAGAGCTTCTAAACAGGTGT 4318  
3314 CTTCAAAGGAGCTCCATCAAACTTGCAGTGTAGTTCTAAAGGGCTTTGTCAACATTTTG 3373  
4319 CTTCAAAGGAGCTCCATCAAACTTGCAGTGTAGTTCTAAAGGGCTTTGTCAACATTTTG 4378  
3374 ACBAACATTGTGCACAAATTTGTTGATTAATTCATTAATCTGTTCCGTTTTCTGATAGG 3433  
4379 ACBAACATTGTGCACAAATTTGTTGATTAATTCATTAATCTGTTCCGTTTTCTGATAGG 4438  
3434 ACBAACAGGAGCTCTTCAATTTCTTGTGAGAGGCTTGTAGAACTCAATTTTGCACAA 3493  
4439 ACBAACAGGAGCTCTTCAATTTCTTGTGAGAGGCTTGTAGAACTCAATTTTGCACAA 4498  
3494 TTCAATACACCTTTCAGAGGAGCTCTTTCATGTTTGGATTCCTTTCATAGGTTGTTCCTT 3553  
4499 TTCAATACACCTTTCAGAGGAGCTCTTTCATGTTTGGATTCCTTTCATAGGTTGTTCCTT 4558  
3554 GGTGTATCTGCGCTTGGCATCTCTTCTTCTAGTGAACCTTTAGGGACTTCATATCCAG 3613

4559 GGTGTATCTGCGCTTGGCATCTCTTCTTCTTCTAGTGACCTTTAGGAGCTTCATATCCAG 4618  
3614 GTTTCTCTCCACCTCTGTCNAACGTCACACCGTACTTGGACATCTAACTAATCCAAATA 3673  
4619 GTTTCTCTCCACCTCTGTCNAACGTCACACCGTACTTGGACATCTAACTAATCCAAATA 4678  
3674 AAATAAGTCAGACATTTCCAGGCTATATCTTCTTCTGGATTTAGCTTCTGCAAGTTCATC 3733  
4679 AAATAAGTCAGACATTTCCAGGCTATATCTTCTTCTGGATTTAGCTTCTGCAAGTTCATC 4738  
3734 AGCTTCTCTCTTAAATTTTATAGGTTTCAACAAACTTCTGCTGCTCAAAATAACCGTTGGTAT 3792  
4739 AGCTTCTCTCTTAAATTTTATAGGTTTCAACAAACTTCTGCTGCTCAAAATAACCGTTGGTAT 4798  
3793 AAGAACCTTCTGAGCATTTCTTACGATCCCAACAGGTGGCTTCCATGGCTCTTAAGAC 3852  
4799 AAGAACCTTCTGAGCATTTCTTACGATCCCAACAGGTGGCTTCCATGGCTCTTAAGAC 4857  
3853 CCTTTGATTGGCCCAAAACAGGAAGTGGTTCGAAGTGAAGAGAACCAACACTTGTGTTT 3912  
4858 CCTTTGATTGGCCCAAAACAGGAAGTGGTTCGAAGTGAAGAGAACCAACACTTGTGTTT 4917  
3913 CAACCAAAATTTCAAGCAGTCTCCATCAAACTTCCAAATTCGAATCCAGCACTTTTGG 3972  
4918 CAACCAAAATTTCAAGCAGTCTCCATCAAACTTCCAAATTCGAATCCAGCACTTTTGG 4977  
3973 TTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGAAGCAGAGATTTGTTAGACTCCAG 4032  
4978 TTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGAAGCAGAGATTTGTTAGACTCCAG 5037  
4033 TTTGTTCTTATAGCTTCGGGAATAGACTTTTGTGACGAGTACACAGGCCCAACAGGT 4092  
5038 TTTGTTCTTATAGCTTCGGGAATAGACTTTTGTGACGAGTACACAGGCCCAACAGGT 5097  
4093 AATTAGAGAGTTCAGCCACCAAAAGTGTGAATAGACATCGGGGGCTCAGTAGTCAAAAG 4152  
5098 AATTAGAGAGTTCAGCCACCAAAAGTGTGAATAGACATCGGGGGCTCAGTAGTCAAAAG 5157  
4153 ACGCCAAACAAATTTCTGACAGGGAACCTTTTGTGACATCTTCAGAAAGTTCGTTATTCAG 4212  
5158 ACGCCAAACAAATTTCTGACAGGGAACCTTTTGTGACATCTTCAGAAAGTTCGTTATTCAG 5217  
4213 TAGTCAATTCGCCAGATCAATTAATGGGATTTATACAGAGCAACAGTGAAGTCAACAT 4272  
5218 TAGTCAATTCGCCAGATCAATTAATGGGATTTATACAGAGCAACAGTGAAGTCAACAT 5277  
4273 CTACCAACTTTTGGCGTCTCAGAAAAAGCATAAACAGTTCCTACTACCGCCATTTAGTGAAC 4332  
5278 CTACCAACTTTTGGCGTCTCAGAAAAAGCATAAACAGTTCCTACTACCGCCATTTAGTGAAC 5337  
4333 TTTTCAAAATCGCCAGTGGAGAGAAAAAGGCAAGCATAGTACATTTAGCGGCAAGG 4392  
5338 TTTTCAAAATCGCCAGTGGAGAGAAAAAGGCAAGCATAGTACATTTAGCGGCAAGG 5397  
4393 ATGCAACTTTTATCAACAGGCTCTATAGATAACCTTAGCGCTGGATCATCTTTTGA 4452  
5398 ATGCAACTTTTATCAACAGGCTCTATAGATAACCTTAGCGCTGGATCATCTTTTGA 5457  
4453 CAACCTTTCTGCAAAATCTAGGTCCAAATCACTTTCATTTGATACCATTTTGTACAACT 4512  
5458 CAACCTTTCTGCAAAATCTAGGTCCAAATCACTTTCATTTGATACCATTTTGTACAACT 5517  
4513 TGAGCAAGTTGTGATCAGCTCTCAAAATTTGCTCTCTGTAACGGATGACTCAACTTTGCA 4572  
5518 TGAGCAAGTTGTGATCAGCTCTCAAAATTTGCTCTCTGTAACGGATGACTCAACTTTGCA 5577  
4573 CATTAACCTGAAGCTCAGTCAATTTAGTGAACCTTTGATCAGGTTTGTGAGTGTGAGCAG 4632  
5578 CATTAACCTGAAGCTCAGTCAATTTAGTGAACCTTTGATCAGGTTTGTGAGTGTGAGCAG 5637  
4633 CATAGGGAACACAGGCTTTTCTTACCAAACTCAAGGAATTTATCAAACTCTGCAACACTTG 4692  
5638 CATAGGGAACACAGGCTTTTCTTACCAAACTCAAGGAATTTATCAAACTCTGCAACACTTG 5697

QY 4693 CGTATGAGGTAGCAAGGAAATGTCATCTTGAAGTCGACAGTGAAGTGTAGTCTTGTAG 4752  
 Db 5698 CGTATGAGGTAGCAAGGAAATGTCATCTTGAAGTCGACAGTGAAGTGTAGTCTTGTAG 5757  
 QY 4753 AAATCTCAAGCCGTATTTTATATCAGTGAAGTCAGTCAATCAGGATCCTCTAGCCG 4812  
 Db 5758 AAATCTCAAGCCGTATTTTATATCAGTGAAGTCAGTCAATCAGGATCCTCTAGCCG 5817  
 QY 4813 GACGATCGTGGCCGACCTGACGTCGCGATCAGCCGCGCCACAGGTGCGGTTGCTGGCG 4872  
 Db 5818 GACGATCGTGGCCGACCTGACGTCGCGATCAGCCGCGCCACAGGTGCGGTTGCTGGCG 5877  
 QY 4873 CCTATATCGCGACATCAACGATGGGGAAGATCGGGCTCGCCATTCGGGCTCATGAGCG 4932  
 Db 5878 CCTATATCGCGACATCAACGATGGGGAAGATCGGGCTCGCCATTCGGGCTCATGAGCG 5937  
 QY 4933 CTGTGTTGCGGTGGGTATGTTGGCAGGCCCGCTGGCGGGGACTGTGTTGGGCGCCATCT 4992  
 Db 5938 CTGTGTTGCGGTGGGTATGTTGGCAGGCCCGCTGGCGGGGACTGTGTTGGGCGCCATCT 5997  
 QY 4993 CTTTCATGCAC----- 5004  
 Db 5998 CTTTCGACCTGCGAGGGGGGGGGAAGCCAGCTTGTGTCTCAAAATCTCTGATGTA 6057  
 QY 5005 ----- 5004  
 Db 6058 CATTGCAACAATAAATAATATCATCATGAACAAATAAACTGCTGTCTACATAAACAG 6117  
 QY 5005 ----- 5004  
 Db 6118 TAATACAGGGGTGTTATGAGCCATATTCAACGGGAACGCTTGTCTCAAGCGCGGATT 6177  
 QY 5005 ----- 5004  
 Db 6178 AAATTCACACATGATGCTGATTTATATGGGTATAAATGGGCTCGCGATATGTGCGGCA 6237  
 QY 5005 ----- 5004  
 Db 6238 ATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTCTGAA 6297  
 QY 5005 ----- 5004  
 Db 6298 ACATGGCAAGGTAGGTTGCCAATGATTTACAGATGAGATGCTCAGACTAAACTGGCT 6357  
 QY 5005 ----- 5004  
 Db 6358 GACGGAATTTATGCTCTTCCGACCATCAAGCAATTTTATCGTACTCCTGATGATGCATG 6417  
 QY 5005 ----- 5004  
 Db 6418 GTTACTCACCACTGCGATCCCCGGGAAACAGCAATTCAGGTATTAGAGAATATCTCTGA 6477  
 QY 5005 ----- 5004  
 Db 6478 TTCAAGTGAANAATGTTGTGATGCGCTGGCAGTGTCTGCGCGGTTGCAATTCGATTCC 6537  
 QY 5005 ----- 5004  
 Db 6538 TGTGTAATTTGCTCTTTTAAACAGCATCGCGTATTTGCTCTGCTCAGGCGCAATCAG 6597  
 QY 5005 ----- 5004  
 Db 6598 AATGAATAACGGTTTGGTTGATGCGAGTGAATTTGATGACGAGCGTAATGCGTGGCTGT 6657  
 QY 5005 ----- 5004  
 Db 6658 TGAAACAGTCTGGAAGAAATGCAATAGCTTTTGCCATTTCTCACCGGATTCTCAGTCTCAC 6717  
 QY 5005 ----- 5004  
 Db 6718 TCATGGTATTTCTCACTTTGATAACCTTATTTTTCAGCGAGGGAATTAATAGTGTGAT 6777

QY 5005 ----- 5004  
 Db 6778 TGATTTGACGAGTCCGAATTCGACAGCCGATACCAGGATCTTCCATCTCTATGAACTG 6837  
 QY 5005 ----- 5004  
 Db 6838 CCTCGGTGAGTTTCTCCTTCATTACAGAAACGGCTTTTCAAAAATATGATATTGATAA 6897  
 QY 5005 ----- 5004  
 Db 6898 TCTCATATGAATAAATTCAGATTTTCAATTCGATGCTCGATGAGTTTCTTAATCAGAAAT 6957  
 QY 5005 ----- 5004  
 Db 6958 GGTAAATGTTGTAACTGCGACAGCATTTACGCTGACTTGAACGGAGCGGCTTTGT 7017  
 QY 5005 ----- 5004  
 Db 7018 TGAATAAATCGAACTTTTCTGAGTTGAAGATCAGATCAGCATCTTCCGACAAACGA 7077  
 QY 5005 ----- 5004  
 Db 7078 GACCGTTCGTTGGCAAGCAAAAGTTCAAATCACAACCTGCTCCACCTACAAAGCT 7137  
 QY 5005 ----- 5004  
 Db 7138 CTATCAACCGTGGTCCCTCACCTTTCTGCTGGATGATGGGGGATTCAGGCTGTGAT 7197  
 QY 5005 ----- 5012  
 Db 7198 GAGTCAGCAACACTTCTTACAGGCGAGACCTCAGCGCCCCCCCCCTTCAGGTCCC 7257  
 QY 5013 GCGGGCGGTGCTCAACGGCTCAACCTATCTATGAGTCTTCTTAATGAGAGTCTG 5072  
 Db 7258 ACGGGCGGTGCTCAACGGCTCAACCTATCTATGAGTCTTCTTAATGAGAGTCTG 7317  
 QY 5073 CATAGGAGAGGTCGAGTATCTATGATGGAAGTATGGGAATGATGATACCCGCTTC 5132  
 Db 7318 CATAGGAGAGGTCGAGTATCTATGATGGAAGTATGGGAATGATGATACCCGCTTC 7377  
 QY 5133 TTCAGTGTCTTCAAGTCTCTATCAGATTTATCCCACTTAAAGCAACCGAGGAGAT 5192  
 Db 7378 TTCAGTGTCTTCAAGTCTCTATCAGATTTATCCCACTTAAAGCAACCGAGGAGAT 7437  
 QY 5193 TTCATGGTAAATTTCTGACTTTTGGTCACTCAGTAGACTCGAACTGTGAGACTATCTCG 5252  
 Db 7438 TTCATGGTAAATTTCTGACTTTTGGTCACTCAGTAGACTCGAACTGTGAGACTATCTCG 7497  
 QY 5253 GTTATCAGCAGCAAAATGCTCTTCTTGGAGACAGTAAATGAAAGTCCCACTTAAAGAAA 5312  
 Db 7498 GTTATCAGCAGCAAAATGCTCTTCTTGGAGACAGTAAATGAAAGTCCCACTTAAAGAAA 7557  
 QY 5313 TCTTCTTATCAGGAAACAACTTCTTGTTCGAACTTTTTCGTCGCTTGAATATAA 5372  
 Db 7558 TCTTCTTATCAGGAAACAACTTCTTGTTCGAACTTTTTCGTCGCTTGAATATAA 7617  
 QY 5373 TCTAGAGTGGATGTCGGGTAGGATGAGCGGCAATGCTTACCTTCTGACCTTCA 5432  
 Db 7618 TCTAGAGTGGATGTCGGGTAGGATGAGCGGCAATGCTTACCTTCTGACCTTCA 7677  
 QY 5433 AGAGGTATGAGGTTTGTAGTACTGATGCAACTTCAAGTTCAGTCAACCGTGTCTTGT 5492  
 Db 7678 AGAGGTATGAGGTTTGTAGTACTGATGCAACTTCAAGTTCAGTCAACCGTGTCTTGT 7737  
 QY 5493 TCAACCATTCGAAATCCAGAGAAATCAAAGTGTGTTGCTACTATTGATCCAAAGCAGT 5552  
 Db 7738 TCAACCATTCGAAATCCAGAGAAATCAAAGTGTGTTGCTACTATTGATCCAAAGCAGT 7797  
 QY 5553 GCGGTCTTGAACCTGACAATAGTGTCTGCTGTTTGGAGTCACTTTTGTATGATTAAT 5612  
 Db 7798 GCGGTCTTGAACCTGACAATAGTGTCTGCTGTTTGGAGTCACTTTTGTATGATTAAT 7857  
 QY 5613 CTAGTCTTTGATCTAATAATCTTTCAGAGCCCAAGCGGATTAATACCCAACTTAAACT 5672

7858 CTAGCTTTGATCTAAATAATCTTGACGAGCCAAAGCGATAAATACCCAAATCTAAACT 7917  
5673 CTTTAAACGTTAAAGACAAAGTATGTCTGCTGTTATTAACCCCAAAATCAGCTCGTA 5732  
7918 CTTTAAACGTTAAAGACAAAGTATGTCTGCTGTTATTAACCCCAAAATCAGCTCGTA 7977  
5733 GTCTGATCTCATCAACTTGAGGGGCACTATCTGTTTATAGAGAAATTTGCGGAGATCG 5792  
7978 GTCTGATCTCATCAACTTGAGGGGCACTATCTGTTTATAGAGAAATTTGCGGAGATCG 8037  
5793 ATATCGAGAAAAGGTAGCTGATTTTAAACGTGAAATTTATCTCAAGATCTCTGCTCG 5852  
8038 ATATCGAGAAAAGGTAGCTGATTTTAAACGTGAAATTTATCTCAAGATCTCTGCTCG 8097  
5853 CGCGTTTCGGTGATGACGGTGAACCTCTGACACATGACAGTCCCGGAGACGFTCACAG 5912  
8098 CGCGTTTCGGTGATGACGGTGAACCTCTGACACATGACAGTCCCGGAGACGFTCACAG 8157  
5913 CTTGCTGTGTAAGCGGATCGGGAGCAGACAAGCCCGTCAGGGCCGTCAGCGGTGTTG 5972  
8158 CTTGCTGTGTAAGCGGATCGGGAGCAGACAAGCCCGTCAGGGCCGTCAGCGGTGTTG 8217  
5973 GCGGTGTCGGGGCCGACCATGACCCAGTCAGAGTGACAGTACAGTACGAGTGTATCTGCT 6032  
8218 GCGGTGTCGGGGCCGACCATGACCCAGTCAGAGTGACAGTACAGTACGAGTGTATCTGCT 8277  
6033 TAACTATGCGGATCAGAGCAGATTTGTAAGAGTGACAGTACAGTACGAGTGTATCTGCT 6092  
8278 TAACTATGCGGATCAGAGCAGATTTGTAAGAGTGACAGTACAGTACGAGTGTATCTGCT 8337  
6093 GCACAGATGCGTAAGAGAAATATCGGATCAGGGCTCTTCGCTTCTCTGCTCAGTGA 6152  
8338 GCACAGATGCGTAAGAGAAATATCGGATCAGGGCTCTTCGCTTCTCTGCTCAGTGA 8397  
6153 CTCGCTGCGCTCGTCTGCTGCTCGGAGCGGTATCAGCTCACTCAAAAGGCGTAAT 6212  
8398 CTCGCTGCGCTCGTCTGCTGCTCGGAGCGGTATCAGCTCACTCAAAAGGCGTAAT 8457  
6213 ACGGTTATCCAGAAATCAGGGATACGAGAAAGAAACATGTGAGCAAAAGGCCAGCA 6272  
8458 ACGGTTATCCAGAAATCAGGGATACGAGAAAGAAACATGTGAGCAAAAGGCCAGCA 8517  
6273 AAAGGCCAGAACCTTAAAGGCGGTTGCTGCTGCGGTTTTCATAGGCTCCGCCCCC 6332  
8518 AAAGGCCAGAACCTTAAAGGCGGTTGCTGCTGCGGTTTTCATAGGCTCCGCCCCC 8577  
6333 TCACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATA 6392  
8578 TCACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATA 8637  
6393 AAGATACAGGGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCC 6452  
8638 AAGATACAGGGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCC 8697  
6453 GCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAAAGCGTGGGCTTCTCATAGCTC 6512  
8698 GCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAAAGCGTGGGCTTCTCATAGCTC 8757  
6513 ACGTGTAGGATCTCAGTTCGGTGTAGTCTGTTCCGCTCCAGCTGGGCTGTGTGACGA 6572  
8758 ACGTGTAGGATCTCAGTTCGGTGTAGTCTGTTCCGCTCCAGCTGGGCTGTGTGACGA 8817  
6573 ACCCCCGTTACGCCCGCTGCGCTTATCCGCTAACCTATCGTCTGAGTCCCAACC 6632  
8818 ACCCCCGTTACGCCCGCTGCGCTTATCCGCTAACCTATCGTCTGAGTCCCAACC 8877  
6633 GGTAAAGACAGCTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATAGCAGCGAG 6692  
8878 GGTAAAGACAGCTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATAGCAGCGAG 8937  
6693 GTATGAGCGGTGCTACAGATTTCTGAAAGTGGTGGCTTAACCTAGCGCTACATAGAG 6752

8938 GTATGAGCGGTCTACAGAGTTCTTGAAGTGGTGCCTTAACCTACGGCTACACTAGAAG 8997  
6753 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAAGTGGTAG 6812  
8998 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAAGTGGTAG 9057  
6813 CTCTTCATCCGGCAAAACAAACACCGCTGCTAGCGGTGGTGTGTTTGTGTTGCAAGCAGCA 6872  
9058 CTCTTCATCCGGCAAAACAAACACCGCTGCTAGCGGTGGTGTGTTTGTGTTGCAAGCAGCA 9117  
6873 GATTACGCGCAGAAAAAAGGATCTCAAGNAGATCTCTTGTATCTTTCACGGGCTCTGA 6932  
9118 GATTACGCGCAGAAAAAAGGATCTCAAGNAGATCTCTTGTATCTTTCACGGGCTCTGA 9177  
6933 CGCTCAGTGAACGAAAACTCAGGTTTAAAGGATTTTGGTCATCAGATTTATCAAAAAGGAT 6992  
9178 CGCTCAGTGAACGAAAACTCAGGTTTAAAGGATTTTGGTCATCAGATTTATCAAAAAGGAT 9237  
6993 CTTACACATGATCTCTTTAAATTAAGATGAGTTTAAATCAATCTAAAGTATATATGA 7052  
9238 CTTACACATGATCTCTTTAAATTAAGATGAGTTTAAATCAATCTAAAGTATATATGA 9297  
7053 GTAAACTTGTCTGACAGTTTACCAATGCTTAACTCAGTGAAGGACCTATCTCAGCGATCTG 7112  
9298 GTAAACTTGTCTGACAGTTTACCAATGCTTAACTCAGTGAAGGACCTATCTCAGCGATCTG 9357  
7113 TCTATTTGTTTCATCCATAGTTCCCTGACTCCCGTCTGTGTAGATAAATACGATACGGGA 7172  
9358 TCTATTTGTTTCATCCATAGTTCCCTGACTCCCGTCTGTGTAGATAAATACGATACGGGA 9417  
7173 GGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCCAGCTCAGCGGCTCC 7232  
9418 GGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCCAGCTCAGCGGCTCC 9477  
7233 AGATTTATCAGCAATAAACCAGCAGCCGGAAGGCGGAGCAGAAAGTGGTCTCTGCAAC 7292  
9478 AGATTTATCAGCAATAAACCAGCAGCCGGAAGGCGGAGCAGAAAGTGGTCTCTGCAAC 9537  
7293 TTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTAGAGTAAGTGTGCGC 7352  
9538 TTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTAGAGTAAGTGTGCGC 9597  
7353 AGTTAATAGTTTCGGCAACGTTGTTGCTGCTGAGGATCGTGGTGTGACGCTCGTC 7412  
9598 AGTTAATAGTTTCGGCAACGTTGTTGCTGCTGAGGATCGTGGTGTGACGCTCGTC 9657  
7413 GTTGTGATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGCGGAGTTACATGATCCCC 7472  
9658 GTTGTGATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGCGGAGTTACATGATCCCC 9717  
7473 CATGTTGTGCAAAAAGCGGTTAGCTCTTCGGTCTCCGATCGTTGTGAGAAAGTAAATT 7532  
9718 CATGTTGTGCAAAAAGCGGTTAGCTCTTCGGTCTCCGATCGTTGTGAGAAAGTAAATT 9777  
7533 GGCGCGAGTGTATCACTCATGTTATGCGACACTGCATAATTTCTTACTGTCTATGCC 7592  
9778 GGCGCGAGTGTATCACTCATGTTATGCGACACTGCATAATTTCTTACTGTCTATGCC 9837  
7593 ATCCGTAAAGATGCTTTTCTGTGATCGGTGAGTACTCAACCAAGTCTATCTGAGAAATAGTG 7652  
9838 ATCCGTAAAGATGCTTTTCTGTGATCGGTGAGTACTCAACCAAGTCTATCTGAGAAATAGTG 9897  
7653 TATGCGCGACCGAGTTGCTTTCGGCGGCTCAACAGCGGATAATACCGCGCCACATAG 7712  
9898 TATGCGCGACCGAGTTGCTTTCGGCGGCTCAACAGCGGATAATACCGCGCCACATAG 9957  
7713 CAGAACTTTAAAGTGTCTCATCTTGGAAAAAGTCTTTCGGGCGGAAAACTCTCAAGGAT 7772  
9958 CAGAACTTTAAAGTGTCTCATCTTGGAAAAAGTCTTTCGGGCGGAAAACTCTCAAGGAT 10017  
7773 CTTACCGCTCTGAGATCCAGTTCCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC 7832  
10018 CTTACCGCTCTGAGATCCAGTTCCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC 10077





QY 1253 -----GTATCATGCCCCCGGAGAGGATAAATGTGGCTTCCCGGTGTC 1299  
Db 1381 ACATCATATTAAAGGATGAAGAAACACCGGAGGACAGGCTGATGACTCAGAGGTGA 1440  
QY 1300 A----- 1300  
Db 1441 AGGACAAGTTGAACGTTTGGCCATCTCGGTGATGAACCACTGGCCAGGAGTGAACCTGC 1500  
QY 1301 ----- 1300  
Db 1501 GGGTACCGAGGGCTGGGACGNAGATGCCACCCTCAGAGGAGTCTCTGCACCTACGAGG 1560  
QY 1301 ----- 1300  
Db 1561 GCGCGCAGTGGACATACCACTGCTGACCGCGACCGCAGCAAGTACGGCATGCTGGCCC 1620  
QY 1301 ----- 1300  
Db 1621 GCCTGGCGGTGAGGCCCGGCTTCGACTGGGTGTACTAGTGTCCAGGCACATATCCACT 1680  
QY 1301 ----- 1300  
Db 1681 GCTCGGTGAAGACAGAGAACTCGGTGGCGGCCAAATCGGGAGGGGTGACCCAGAGGGC 1740  
QY 1301 ----- 1300  
Db 1741 CCACAAATCAAGCCCTGTCTCCATGCAAAATGCCAGCACCTAACCTTTGGGTGGACCAT 1800  
QY 1301 ----- 1300  
Db 1801 CCGTCTTCATCTTCCCTCCAAAGATCAAGGATGTACTCATGATCTCCCTGAGCCCCATAG 1860  
QY 1301 ----- 1300  
Db 1861 TCACATGTGTGGTGGATGTGAGCGAGGATGACCCAGATGTCACAGTACGTGGTTG 1920  
QY 1301 ----- 1300  
Db 1921 TGAACAACGTGGAAGTACACACAGCTCAGACACACAAACCCATAGAGAGATTACCAAGTA 1980  
QY 1301 ----- 1300  
Db 1981 CACTTCGGTGTGTCAGTGCCCTCCCTCCAGCACAGGACTGGATGAGTGGCAAGGAGT 2040  
QY 1301 ----- 1300  
Db 2041 TCAAAATGCAAGTCAACAAACAAGACCTCCAGCGCCCATCGAGAGAACCATCTCAAAAC 2100  
QY 1301 ----- 1300  
Db 2101 CCAAAGGTTCAGTAAGACTCCACAGGTATATGTCTTCCCTCCACCAGAGAGATGA 2160  
QY 1301 -----CCGCCACGAGTGCACGGAGAGAGGTGTGTTTGTATGACAGTGTCC 1348  
Db 2161 CTAAGAAAACAGGTCACTCTGACCTGCTGATGAGTACAGACTTCATGCTGAGACATTTACG 2220  
QY 1349 GGG----- 1351  
Db 2221 TGGAGTGGACCAACAAACGGGAAAAACAGAGCTTAACTACAGAACACTGAAACAGTCTCTGG 2280  
QY 1352 -----GATTCGCGTGGTGTCTCCACCCCATGGCCATCGCAACACTCAAG 1396  
Db 2281 ACTCTGATGGTCTTACTTCTATGTACAGCAAGCTGAGAGTGGAAAGAAAGTGGGTGG 2340  
QY 1397 AAGAGAAATGTCCTCTTAACTAGTGG----- 1423  
Db 2341 AAAGAAAATAGTACTCTCTGTTTCAGTGGTCCAGAGGGTCTGCACAAATCACACAGACTA 2400  
QY 1424 -----CGTGAATTC 1433  
Db 2401 AGAGCTTCTCCGGACTCCGGGTAAATGAGCTCAGATCGAATTCATCGATCTCCATCTC 2460

QY 1434 CTTAGGCGCGCCGCAATTAATTCGCTTAGACATGACTGTTCCTCAGTTCAGTTGGGC 1493  
Db 2461 CAATCCGCGCGCCGCAATTAATTCGCTTAGACATGACTGTTCCTCAGTTCAGTTGGGC 2520  
QY 1494 ACTTCGAGAGAGACCGGTCTTCTAGATCTCAATCAAGAGGATGTGAGATGCCATTTCG 1553  
Db 2521 ACTTCGAGAGAGACCGGTCTTCTAGATCTCAATCAAGAGGATGTGAGATGCCATTTCG 2580  
QY 1554 CTGAGAGATGACAGCTTCATTTTGTATCTTTTATTTTAACTATATAGTATAGGAT 1613  
Db 2581 CTGAGAGATGACAGCTTCATTTTGTATCTTTTATTTTAACTATATAGTATAGGAT 2640  
QY 1614 TTTTTCATTTTCTTCTCTGACAGCTGCTCTCTGATGAGCTATCTCCAGC 1673  
Db 2641 TTTTTCATTTTCTTCTCTGACAGCTGCTCTCTGATGAGCTATCTCCAGC 2700  
QY 1674 TGATGAATATCTTGTGTAGGGTTTGGGAAATCAATTCAGATTTGATGTTTCTTGT 1733  
Db 2701 TGATGAATATCTTGTGTAGGGTTTGGGAAATCAATTCAGATTTGATGTTTCTTGT 2760  
QY 1734 ATTTCACCTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTCGTTTGTGCAAGCTTA 1793  
Db 2761 ATTTCACCTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTCGTTTGTGCAAGCTTA 2820  
QY 1794 TCGATAGCTTTAATCGGTAGTTTATCACAGTTAAATGCTAAACAGTTCAGGACCGT 1853  
Db 2821 TCGATAGCTTTAATCGGTAGTTTATCACAGTTAAATGCTAAACAGTTCAGGACCGT 2880  
QY 1854 GTATGAATCTAAATGCGCTCATCTCGGCAACCGTCAACCTGGATGCTGTAG 1913  
Db 2881 GTATGAATCTAAATGCGCTCATCTCGGCAACCGTCAACCTGGATGCTGTAG 2940  
QY 1914 GCATAGCTTGTATGCGGTACTGCGGGCTCTTGGGGATATCGTCCATTCGACA 1973  
Db 2941 GCATAGCTTGTATGCGGTACTGCGGGCTCTTGGGGATATCGTCCATTCGACA 3000  
QY 1974 GCATCGCAGTCACTATGCGTCTAGCGTATATGCGTTGATGCAATTTCTATGCG 2033  
Db 3001 GCATCGCAGTCACTATGCGTCTAGCGTATATGCGTTGATGCAATTTCTATGCG 3060  
QY 2034 CACCGTTCTCGAGACACTGTCCGACCGTTTGGCGCGCCCGAGTCTCTGCTGCTTCG 2093  
Db 3061 CACCGTTCTCGAGACACTGTCCGACCGTTTGGCGCGCCCGAGTCTCTGCTGCTTCG 3120  
QY 2094 TACTTCGAGCCTATCGACTACGCGATCATGCGGACCAACCCGCTCTGTGATCTATC 2153  
Db 3121 TACTTCGAGCCTATCGACTACGCGATCATGCGGACCAACCCGCTCTGTGATCTATC 3180  
QY 2154 GAATCTAAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATCGA 2213  
Db 3181 GAATCTAAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATCGA 3240  
QY 2214 ACCTTAACAGCATGCGGTAGCATCTAGACCTTCAACGAGCAGCAGATCCATCTGCT 2273  
Db 3241 ACCTTAACAGCATGCGGTAGCATCTAGACCTTCAACGAGCAGCAGATCCATCTGCT 3300  
QY 2274 TGGCCAAATATGTTTCAGTCCCTCAGGATTAACGTTTGTGAAGTGAATCTCTGGAAG 2333  
Db 3301 TGGCCAAATATGTTTCAGTCCCTCAGGATTAACGTTTGTGAAGTGAATCTCTGGAAG 3360  
QY 2334 GTTTCAGTGTAACTCCGCTGATTAAGCGGCAATTCGTAAGTGGGAAAGTGGTTG 2393  
Db 3361 GTTTCAGTGTAACTCCGCTGATTAAGCGGCAATTCGTAAGTGGGAAAGTGGTTG 3420  
QY 2394 GTACCGGAGGATAATCTCCAACTCTCTGGAGATGAGCACAACAAACACAGATCCA 2453  
Db 3421 GTACCGGAGGATAATCTCCAACTCTCTGGAGATGAGCACAACAAACACAGATCCA 3480  
QY 2454 GCGTGTGTATCTGATCAACATAAGAGAGCAATCTCGATTTGAGGATCAAGTGTTC 2513  
Db 3481 GCGTGTGTATCTGATCAACATAAGAGAGCAATCTCGATTTGAGGATCAAGTGTTC 3540  
QY 2514 GGAGCGTACTGATTGGACATTTCCAAAGCCTGCTCTAGTTGCAACCGATAGGGTTGA 2573

Db	3541	GGAGCGTACTGATTGGACATTTCCAAAGCCTGCTCGTAGTGTGCACCCGATAGGGTTGTA	3600
QY	2574	GAGTGTGCATACACTTGGCTGACAAATTTCAACCTTGGCAACTGCACAGCTTGTTGTGA	2633
Db	3601	GAGTGTGCAATACACTTGGCTGACAAATTTCAACCTTGGCAACTGCACAGCTTGTTGTGA	3660
QY	2634	ACAGCATCTTCAATTTCTGGCAAGCTCTTGCTGTCTCATATCGACAGCCAAACAGAAATCACC	2693
Db	3661	ACAGCATCTTCAATTTCTGGCAAGCTCTTGCTGTCTCATATCGACAGCCAAACAGAAATCACC	3720
QY	2694	TGGGAATCAATATCAATGTTTCAAGTTTGTGACAGAGAGTCTGAGGCAACGAATCTGATCA	2753
Db	3721	TGGGAATCAATATCAATGTTTCAAGTTTGTGAG - CAGAGAGTCTGAGGCAACGAATCTGATCA	3779
QY	2754	CGGTATTTATCAGCAATAACTAGAACTTCAGAGAGCCACAGGCGATGTCAATACTACAC	2813
Db	3780	CGGTATTTATCAGCAATAACTAGAACTTCAGAGAGCCACAGGCGATGTCAATACTACAC	3839
QY	2814	AGGCTGATGTGTCATTTTGAACAATCATCTTGGCAGCAGTAACGAATCTGGTTTCTGTGA	2873
Db	3840	AGGCTGATGTGTCATTTTGAACAATCATCTTGGCAGCAGTAACGAATCTGGTTTCTGTGA	3899
QY	2874	CCAAATATTTGTCCACACTTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGTACTGCC	2933
Db	3900	CCAAATATTTGTCCACACTTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGTACTGCC	3959
QY	2934	TGGCGGCTCTCTGTAGCAGATACACTTAGCACCAACCTTGTGGGCAACGTTAGATGACT	2993
Db	3960	TGGCGGCTCTCTGTAGCAGATACACTTAGCACCAACCTTGTGGGCAACGTTAGATGACT	4019
QY	2994	TCTGGGTAGGGTACCATCTTCTTAGTGTGGAGATGCAAAACAATTTCTTTGCAACCA	3053
Db	4020	TCTGGGTAGGGTACCATCTTCTTAGTGTGGAGATGCAAAACAATTTCTTTGCAACCA	4079
QY	3054	GCAACTTTGGCAGGAAACCCAGCATCAGGGAAAGTGAAGGCGAAATTCGGGTTCCACCA	3113
Db	4080	GCAACTTTGGCAGGAAACCCAGCATCAGGGAAAGTGAAGGCGAAATTCGGGTTCCACCA	4139
QY	3114	GGAATATAGAGGCCAATTTCTCAATAGTCTTGCATAAAGCAGACGACATACACGAGG	3173
Db	4140	GGAATATAGAGGCCAATTTCTCAATAGTCTTGCATAAAGCAGACGACATACACGAGG	4199
QY	3174	CAAGTCTCAACTTGCACAGTCTCGTTAGTTAGCTTCATGGAAATTTCTTGACGTTATCT	3233
Db	4200	CAAGTCTCAACTTGCACAGTCTCGTTAGTTAGCTTCATGGAAATTTCTTGACGTTATCT	4259
QY	3234	ATAGAGAGATCAATGGTCTCTTAACGTTATCTGGCAATTTGCATAAGTTCCTCTGGGAAA	3293
Db	4260	ATAGAGAGATCAATGGTCTCTTAACGTTATCTGGCAATTTGCATAAGTTCCTCTGGGAAA	4319
QY	3294	GGAGCTCTTAACACAGGTGCTTCCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAA	3353
Db	4320	GGAGCTCTTAACACAGGTGCTTCCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAA	4379
QY	3354	AGGCTTTGTCCACATTTTGACGAACATTTGTCGAATTTGGTTTGAATAATTCATTAATC	3413
Db	4380	AGGCTTTGTCCACATTTTGACGAACATTTGTCGAACATTTGGTTTGAATAATTCATTAATC	4439
QY	3414	TGTTCCGTTTCTGGATAGGACGAGGCGCATCTTCAATTTCTTGTGAGGAGGCCCTTA	3473
Db	4440	TGTTCCGTTTCTGGATAGGACGAGGCGCATCTTCAATTTCTTGTGAGGAGGCCCTTA	4499
QY	3474	GAACAGTCAATTTTGCACAAATTCATATACGACTTCAGAAGGCACTCTTTAGGTTTCGAT	3533
Db	4500	GAACAGTCAATTTTGCACAAATTCATATACGACTTCAGAAGGCACTCTTTAGGTTTCGAT	4559
QY	3534	TCCTTTTAGTGTGTTCTGGTGTATCTGGCTTGGCATCTCTTTCTCTTAGTGACCC	3593
Db	4560	TCCTTTTAGTGTGTTCTGGTGTATCTGGCTTGGCATCTCTTTCTCTTAGTGACCC	4619
QY	3594	TTTAGGACATTCATATCAGGTTTCTCTCCACCTCGTCCAAAGTCACACCGTACTCTTGCA	3653

Db	4620	TTTAGGAGCTTCATATCCAGGTTTCTCTCCACCTCGTCCAAAGTGCACACGGTACTTGGCA	4675
Qy	3654	CATCTAACTAATGCAAAATAAAATAAGTCAGACATTTCCAGGCTATATCTTCCTTTGGAT	3713
Db	4680	CACTAACTAAATGCAAAATAAAATAAGTCAGACATTTCCAGGCTATATCTTCCTTTGGAT	4739
Qy	3714	TTAGCTTTCTGCAAGTTTCATCAGCTTCCTCCCTTAATTTTAGCGTTC-AACAAAACCTTTCCGTC	3772
Db	4740	TTAGCTTTCTGCAAGTTTCATCAGCTTCCTCCCTTAATTTTAGCGTTCAAACAAAACCTTTCCGTC	4799
Qy	3773	GTCAAAATAACCGTTTGGTATAGAAACCTTCTGGAGCATTTGCTCTTACGATCCCAACAGGT	3832
Db	4800	GTCAAAATAACCGTTTGGTATAGAAACCTTCTGGAGCATTTGCTCTTACGATCCCAACAGGT	4859
Qy	3833	GGCTTCGATGGCTCTTAAGACCCCTTGATTTGGCCAAAACAGGAAGTGGTTCACAGTGACA	3892
Db	4860	-GCTTCGATGGCTCTTAAGACCCCTTGATTTGGCCAAAACAGGAAGTGGTTCACAGTGACA	4918
Qy	3893	GAACCAACACCTGTTTGTCTCAACCAAAATTTCAAGCAGTCTCCATCAAAATCCAAATTC	3952
Db	4919	GAACCAACACCTGTTTGTCTCAACCAAAATTTCAAGCAGTCTCCATCAAAATCCAAATTC	4978
Qy	3953	GATACCCAGCAACTTTTGAGTTGCTCCAGATGAGCACCTTTATACCAAAACCGTGACG	4012
Db	4979	GATACCCAGCAACTTTTGAGTTGCTCCAGATGAGCACCTTTATACCAAAACCGTGACG	5038
Qy	4013	ACCAGATTTGGTAGCTCCAGTTTGCTCTTATAGCCTCCGGAATAGACTTTTGTGACGA	4072
Db	5039	ACCAGATTTGGTAGCTCCAGTTTGCTCTTATAGCCTCCGGAATAGACTTTTGTGACGA	5098
Qy	4073	GTACACAGCCCAACGAGTAATTTAGAAGAGTCAGCCACCAAGTAGTAGTAATAGACATC	4132
Db	5099	GTACACAGCCCAACGAGTAATTTAGAAGAGTCAGCCACCAAGTAGTAGTAATAGACATC	5158
Qy	4133	GGGGCGGTGAGTAGTCAAGACGCCAACAAAATTTCACTACAGGGAACCTTTTGTGACATC	4192
Db	5159	GGGGCGGTGAGTAGTCAAGACGCCAACAAAATTTCACTACAGGGAACCTTTTGTGACATC	5218
Qy	4193	TTCAGAAAGTTTCGTAATTCAGTAGTCAATTTGCCGAGCATCAATTAATGGGATTATACCAGA	4252
Db	5219	TTCAGAAAGTTTCGTAATTCAGTAGTCAATTTGCCGAGCATCAATTAATGGGATTATACCAGA	5278
Qy	4253	AGCAACAGTGGGAAGTCACATCTACAACTTTGGGTCTCAGAAAAGCATTAACAGTTCT	4312
Db	5279	AGCAACAGTGGGAAGTCACATCTACAACTTTGGGTCTCAGAAAAGCATTAACAGTTCT	5338
Qy	4313	ACTACCGCATTTAGTGAACCTTTTCAAATCGCCAGTGGAGAAAGAAAAGGCACAGCGAT	4372
Db	5339	ACTACCGCATTTAGTGAACCTTTTCAAATCGCCAGTGGAGAAAGAAAAGGCACAGCGAT	5398
Qy	4373	ACTAGCATTAGCCGGCAAGGATCAACTTTTATCAACACAGGTCCTATAGATATACCCTAGC	4432
Db	5399	ACTAGCATTAGCCGGCAAGGATCAACTTTTATCAACACAGGTCCTATAGATATACCCTAGC	5458
Qy	4433	GCCTGGATCATCCTTTGTGACAACTCTTCTCGCCAAATCTAGTTCGAAAATCACTTCATT	4492
Db	5459	GCCTGGATCATCCTTTGTGACAACTCTTCTCGCCAAATCTAGTTCGAAAATCACTTCATT	5518
Qy	4493	GATACCATTTATTGTACAACCTTGACGAAGTTGTTCGATCAGCTCCTCAAATTTGGTCTCTGT	4552
Db	5519	GATACCATTTATTGTACAACCTTGACGAAGTTGTTCGATCAGCTCCTCAAATTTGGTCTCTGT	5578
Qy	4553	AACGGATGACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTTAGTGAACTTGATCAG	4612
Db	5579	AACGGATGACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTTAGTGAACTTGATCAG	5638
Qy	4613	GTGTGTGAGCTGGTCAGCAGCATAGGGAACACAGGCTTTTCTTACCAAAATCAAGGAATT	4672
Db	5639	GTGTGTGAGCTGGTCAGCAGCATAGGGAACACAGGCTTTTCTTACCAAAATCAAGGAATT	5698
Qy	4673	ATCAAACTCTGGAACACTTCGCTATGCGAGTAGCAGGGAATGTCTACTTCTGAAGTCGG	4732
Db	5699	ATCAAACTCTGGAACACTTCGCTATGCGAGTAGCAGGGAATGTCTACTTCTGAAGTCGG	5758

QY 4733 ACAGTGAGTGTAGTCTTGAGAAATCTGAAAGCGGTATTTTATATATAGTCAAGTCA 4792  
| | | | |  
Db 5759 ACAGTGAGTGTAGTCTTGAGAAATCTGAAAGCGGTATTTTATATAGTCAAGTCA 5818  
| | | | |  
QY 4793 TCAGGAGATCCTCTACGCCGACGATCGTGGCCGACCTGCAGGTCGSCATCACCGCGC 4852  
| | | | |  
Db 5819 TCAGGAGATCCTCTACGCCGACGATCGTGGCCGACCTGCAGGTCGSCATCACCGCGC 5878  
| | | | |  
QY 4853 CACAGTGCGGTGTGCGGCTATATCCCGACATCACCGATGGGAAGATCGGGCTCG 4912  
| | | | |  
Db 5879 CACAGTGCGGTGTGCGGCTATATCCCGACATCACCGATGGGAAGATCGGGCTCG 5938  
| | | | |  
QY 4913 CCATTCGGGCTCATGCGCTGTGTTCCGGGTGGTATGTCAGCCCGCTGGCGG 4972  
| | | | |  
Db 5939 CCATTCGGGCTCATGCGCTGTGTTCCGGGTGGTATGTCAGCCCGCTGGCGG 5998  
| | | | |  
QY 4973 GGGACTGTGGCGGCATCTCCTTCGATGCAC 5004  
| | | | |  
Db 5999 GGGACTGTGGCGGCATCTCCTTCGACCTGCAGGGGGGGGAAAGCCAGTTGTG 6058  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6059 TCTCAAAATCTGTATGTATCATTCACAAGATAAAAATATATCATCATGAACAATAAA 6118  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6119 CTGTCTGTATACATAAAACAGTAATACAAGGGGTGTATGAGCCATATTCACAGGAAACG 6178  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6179 TCTTGCTCAAGGCCCGGATTAATTCACATGATGCTGATTTATATGGGTATAAATGG 6238  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6239 GCTCCGATATGTCCGGCAATCAGTCCGACAAATCTATCGATTGTATGGGAAGCCCGAT 6298  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6299 GGGCAGAGTTGTTCTGAAACATGGCAAGTAGCGTTGCCAATGTTTACAGATCAG 6358  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6359 ATGTCAGACTAACTGGCTGACCGAATTTATGCTCTTCGCAATCAAGCATTTTATC 6418  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6419 CGTACTCCTGATGATGCTGTTACTCACACTCGATCCCGGGGAAACAGCATTCAG 6478  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6479 GTATTAGAAGATATCCTGATTCAGGTGAAATAATGTTGATGCGCTGGCAGTGTCTCG 6538  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6539 GCCTGGTGCATTCGATTCCTGTTGTAATGTTGCTCTTTTAAACAGGATCGCGTATTTCTG 6598  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6599 CTCGCTCAGCGCAATCAGAAATGAATAACGGTTGTTGATGCGAGTATTTTATGATGAC 6658  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6659 GAGCGTAATGGTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTCCCATTC 6718  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6719 TCACCGGATTCAGTCTCACTCATGTTCTCACTTTGATAACCTTATTTTTCAGCAG 6778  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6779 GGGAAATTAATAGTGTATGATGTTGGACGAGTCGGAATCGCAGACCGGATACCAAGAT 6838  
| | | | |

QY 5005 5004  
| | | | |  
Db 6839 CTTGCCATCCTATGGAACCTCCTCGGTGAGTTTCTCTTCAATTACAGAAACGGTTTTT 6898  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6899 CAAAAATATGGTATTGATATCTGATATGAATAAATTCAGATTTTCATTTGATCTCGAT 6958  
| | | | |  
QY 5005 5004  
| | | | |  
Db 6959 GAGTTTTTCTAATCAGAAATTTGTTAATTTGTTGTAACACTGGCAGACCATTTACGCTGACT 7018  
| | | | |  
QY 5005 5004  
| | | | |  
Db 7019 TGACGGACGGCGGCTTTGTTGAATAAATCGAACTTTTCTGAGTTGAAGGATCAGATCA 7078  
| | | | |  
QY 5005 5004  
| | | | |  
Db 7079 CGCATCTTCCCGACAAACGACAGCCGTTCCGTGGCAAAAGTTCAAAATCACCACACT 7138  
| | | | |  
QY 5005 5004  
| | | | |  
Db 7139 GTTCCACTACAAAGAGCTCTCATCAACCGTGGTCCCTCACTTTCTGGCTGGATGATG 7198  
| | | | |  
QY 5005 5004  
| | | | |  
Db 7199 GGGCGATTGAGGCTGTATGATGATCAGCAACACCTTTCTTTCAGAGGACGACCTCAGCGCC 7258  
| | | | |  
QY 5005 5004  
| | | | |  
Db 7259 CCCCCCCTGTCAGGTCCACGGCGGCGGTGCTCAACGGCCTCAACCTACTACTCGGGCT 7318  
| | | | |  
QY 5053 GCTTCTTAATGTCAGGAGTGCATAAGGAGAGCGTGCAGTATCTATGATTTGGAAGTATGG 5112  
| | | | |  
Db 7319 GCTTCTTAATGTCAGGAGTGCATAAGGAGAGCGTGCAGTATCTATGATTTGGAAGTATGG 7378  
| | | | |  
QY 5113 GAATGTGATACCCGATCTTTCAGTGTCTTGAGGTCTCCTATCAGATATGCCCCAACTA 5172  
| | | | |  
Db 7379 GAATGTGATACCCGATCTTTCAGTGTCTTGAGGTCTCCTATCAGATATGCCCCAACTA 7438  
| | | | |  
QY 5173 AGCAACCGGAGGAGGAGATTTTCATGTTAAATTTCTCTGACTTTTGGTCTCATCAGTACT 5232  
| | | | |  
Db 7439 AGCAACCGGAGGAGGAGATTTTCATGTTAAATTTCTCTGACTTTTGGTCTCATCAGTACT 7498  
| | | | |  
QY 5233 CGAACTGTGAGACTATCTCGGTTTATGACAGCAGAAATGTCTTCTTGGAGACAGTAAATG 5292  
| | | | |  
Db 7499 CGAACTGTGAGACTATCTCGGTTTATGACAGCAGAAATGTCTTCTTGGAGACAGTAAATG 7558  
| | | | |  
QY 5293 AGTCCCAACCAATAAAGAAATCTTGTATCAGGAACAACTTCTTGTTCGAACCTTTT 5352  
| | | | |  
Db 7559 AGTCCCAACCAATAAAGAAATCTTGTATCAGGAACAACTTCTTGTTCGAACCTTTT 7618  
| | | | |  
QY 5353 CGGTGCTTGAACCTATAAAATGTAGAGTGGATATGTCGGGTAGGAATGGAGCGGCAAT 5412  
| | | | |  
Db 7619 CGGTGCTTGAACCTATAAATGTAGAGTGGATATGTCGGGTAGGAATGGAGCGGCAAT 7678  
| | | | |  
QY 5413 GCTTACCTTTCGACCTTCAAGAGGTATGTAGGTTTGTAGATPACTGATGCCAATTCAG 5472  
| | | | |  
Db 7679 GCTTACCTTTCGACCTTCAAGAGGTATGTAGGTTTGTAGATPACTGATGCCAATTCAG 7738  
| | | | |  
QY 5473 TGACAAAGTGTCTATTTCTGTTCCAAATCCGATCCGAAATCCAGAGAAATCAAGTTGTTGTC 5532  
| | | | |  
Db 7739 TGACAAAGTGTCTATTTCTGTTCCAAATCCGATCCGAAATCCAGAGAAATCAAGTTGTTGTC 7798  
| | | | |  
QY 5533 TACTATTGATCCAAAGCCAGTGGCGTCTTGAAACTGCAATAGTGTCTCGTGTGTTGAGG 5592  
| | | | |  
Db 7799 TACTATTGATCCAAAGCCAGTGGCGTCTTGAAACTGCAATAGTGTCTCGTGTGTTGAGG 7858  
| | | | |  
QY 5593 TCATCTTTGATGAATAAATCTAGTCTTTGATCTAAATAATCTTTCAGCGACCAAGCGAT 5652  
| | | | |  
Db 7859 TCATCTTTGATGAATAAATCTAGTCTTTGATCTAAATAATCTTTCAGCGACCAAGCGAT 7918  
| | | | |  
QY 5653 AAATACCCAAATCTAAATACTTTTAAACGTTTAAAGGACAAATGATGTCTGCTCTATT 5712  
| | | | |

Db 7919 AATATCCCAATCTAAACTCTTTTAAAGCTTAAAGGCAAGTATGCTGCTGTATT 7978  
Qy 5713 AAACCCCAATCAGCTGCTAGTCTGATCCCTCATCAACTTGAGGGGCACTATCTTGTTTAA 5772  
Db 7979 AAACCCCAATCAGCTGCTAGTCTGATCCCTCATCAACTTGAGGGGCACTATCTTGTTTAA 8038  
Qy 5773 GAGAAATTTGGGAGATCGCATATCGAAGAAAGGTACGCTGATTTTAAACGTGAATTT 5832  
Db 8039 GAGAAATTTGGGAGATCGCATATCGAAGAAAGGTACGCTGATTTTAAACGTGAATTT 8098  
Qy 5833 ATCTCAAGATCTCTGCTCGCGCTTTTCGGTGTAGCGGTGAATACTCTGACACATGCA 5892  
Db 8099 ATCTCAAGATCTCTGCTCGCGCTTTTCGGTGTAGCGGTGAATACTCTGACACATGCA 8158  
Qy 5993 GCTCCCGAGACGCTCAGCTTGTCTGTAAGCGGATGCCGGAGCAGACCAAGCCGTCA 5952  
Db 8159 GCTCCCGAGACGCTCAGCTTGTCTGTAAGCGGATGCCGGAGCAGACCAAGCCGTCA 8218  
Qy 5953 GGGCGCTCAGCGGCTGTTGGCGGTGTCCGGGCGCAGCCATGACCCAGTCAAGTACGCA 6012  
Db 8219 GGGCGCTCAGCGGCTGTTGGCGGTGTCCGGGCGCAGCCATGACCCAGTCAAGTACGCA 8278  
Qy 6013 TAGCGAGTGTATCTGGCTTAATCTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCAC 6072  
Db 8279 TAGCGAGTGTATCTGGCTTAATCTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCAC 8338  
Qy 6073 CATATGGGTGTGAATACCCACAGATCGTGAAGGAGAAATACCCATCAGCGGTCT 6132  
Db 8339 CATATGGGTGTGAATACCCACAGATCGTGAAGGAGAAATACCCATCAGCGGTCT 8398  
Qy 6133 TCCGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 6192  
Db 8399 TCCGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 8458  
Qy 6193 GCTCACTAAAGGGGTGAATACGGTTATCCACAGATCAGGGGATACGAGGAAGAAC 6252  
Db 8459 GCTCACTAAAGGGGTGAATACGGTTATCCACAGATCAGGGGATACGAGGAAGAAC 8518  
Qy 6253 ATGTGAGCAAAAGGCCAGCAAAAGCCAGAAACCGTAAAGGCGCGGTTGCTGGCGTTT 6312  
Db 8519 ATGTGAGCAAAAGGCCAGCAAAAGCCAGAAACCGTAAAGGCGCGGTTGCTGGCGTTT 8578  
Qy 6313 TTCCATAGGCTCCGCCCCCTTGACAGCATCAAAAATCGACCTCAAGTCAGAGTGG 6372  
Db 8579 TTCCATAGGCTCCGCCCCCTTGACAGCATCAAAAATCGACCTCAAGTCAGAGTGG 8638  
Qy 6373 CGAAACCGCAGGACTATAAGATACAGCGGTTTCCCTCGGAAGCTCCCTCGTCGC 6432  
Db 8639 CGAAACCGCAGGACTATAAGATACAGCGGTTTCCCTCGGAAGCTCCCTCGTCGC 8698  
Qy 6433 TCTCCTGTTCCGACCTCCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGC 6492  
Db 8699 TCTCCTGTTCCGACCTCCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGC 8758  
Qy 6493 GTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGGTAGGTGCTTCGCTCC 6552  
Db 8759 GTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGGTAGGTGCTTCGCTCC 8818  
Qy 6553 AAGCTGGGTGTGTGACAGAAACCCCGCTTACGCGGACCGCTCGGCTTTATCCGGTAAC 6612  
Db 8819 AAGCTGGGTGTGTGACAGAAACCCCGCTTACGCGGACCGCTCGGCTTTATCCGGTAAC 8878  
Qy 6613 TATCCTTTGAGTCCAAACCGGTAAAGACAGCACTATTCGCCACTGGCAGGACCACTGCT 6672  
Db 8879 TATCCTTTGAGTCCAAACCGGTAAAGACAGCACTATTCGCCACTGGCAGGACCACTGCT 8938  
Qy 6673 AACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTTGAAGTGTGGCT 6732  
Db 8939 AACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTTGAAGTGTGGCT 8998  
Qy 6733 AACTACGGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACC 9058  
Qy 6793 TTCCGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGT 6852  
Db 9059 TTCCGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGT 9118  
Qy 6853 TTTTGTGTTCCAAAGCAGCAGATTACCGGAGAAAAAGGATCTCAAGAGATCTCTTTG 6912  
Db 9119 TTTTGTGTTCCAAAGCAGCAGATTACCGGAGAAAAAGGATCTCAAGAGATCTCTTTG 9178  
Qy 6913 ATCTTTTCTACGGGCTGTGAGCTCAGTGAAGCAAAACTCACGTTAAGGATTTTGGTC 6972  
Db 9179 ATCTTTTCTACGGGCTGTGAGCTCAGTGAAGCAAAACTCACGTTAAGGATTTTGGTC 9238  
Qy 6973 ATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAATTTAAATTTAA 7032  
Db 9239 ATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAATTTAAATTTAA 9298  
Qy 7033 TCAATCTAAAAGTATATAGTAAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 7092  
Db 9299 TCAATCTAAAAGTATATAGTAAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 9358  
Qy 7093 GCACCTATCTCAGCGATCTGTCTATTTTCGTTTCACTCCATAGTGTGCTGACTCCCGT 7152  
Db 9359 GCACCTATCTCAGCGATCTGTCTATTTTCGTTTCACTCCATAGTGTGCTGACTCCCGT 9418  
Qy 7153 TAGATAAATAAGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 7212  
Db 9419 TAGATAAATAAGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 9478  
Qy 7213 GACCAACGCTCAGCGCTCCAGATTTATCAGCAATATAACCAAGCCAGCGGAGCCGAG 7272  
Db 9479 GACCAACGCTCAGCGCTCCAGATTTATCAGCAATATAACCAAGCCAGCGGAGCCGAG 9538  
Qy 7273 CGCAGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAA 7332  
Db 9539 CGCAGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAA 9598  
Qy 7333 GCTAGATTAAGTGTGCTCCAGTAAATAGTTTGGCAACGTTTGTCCATTTGCTGCAGGC 7392  
Db 9599 GCTAGATTAAGTGTGCTCCAGTAAATAGTTTGGCAACGTTTGTCCATTTGCTGCAGGC 9658  
Qy 7393 ATCTGTGTGTCAAGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGTTCCCAAGATCA 7452  
Db 9659 ATCTGTGTGTCAAGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGTTCCCAAGATCA 9718  
Qy 7453 AGCGAGTTATATGATCCCCATGTTGTGCAAAAAGCGGTAGTCTCTCGTCCCTCCG 7512  
Db 9719 AGCGAGTTATATGATCCCCATGTTGTGCAAAAAGCGGTAGTCTCTCGTCCCTCCG 9778  
Qy 7513 ATCTGTGTGTCAGAAATAGTTGGCCGAGTTATCACTCATGTTTATGGCAGCACTGCAT 7572  
Db 9779 ATCTGTGTGTCAGAAATAGTTGGCCGAGTTATCACTCATGTTTATGGCAGCACTGCAT 9838  
Qy 7573 AATCTCTTATCTGATCCGATCCGATAGATGCTTTCTGTGACTGTGTGACTCAACC 9898  
Db 9839 AATCTCTTATCTGATCCGATCCGATAGATGCTTTCTGTGACTGTGTGACTCAACC 9958  
Qy 7633 AAGTCAATCTGAGATAGTGTATGCGGAGCCAGTGTGCTCTTGGCCGCGCTCAACACGG 7692  
Db 9899 AAGTCAATCTGAGATAGTGTATGCGGAGCCAGTGTGCTCTTGGCCGCGCTCAACACGG 9958  
Qy 7693 GATAATACCGCCCATAGCAGAACTTTAAAGTGTCTCATCTTTGGAAGAAAGTCTTTCG 7752  
Db 9959 GATAATACCGCCCATAGCAGAACTTTAAAGTGTCTCATCTTTGGAAGAAAGTCTTTCG 10018  
Qy 7753 GGGCGAAATCTCTAAGGATCTTACCGTGTGTGAGATCCAGTTCAGTGTAAACCACTCGT 7812  
Db 10019 GGGCGAAATCTCTAAGGATCTTACCGTGTGTGAGATCCAGTTCAGTGTAAACCACTCGT 10078  
Qy 7813 GCACCCAACTGATCTTACGATCTTTTACTTTTCCACGCGTTTCTGGGTGAGCAAAAACA 7872  
Db 10079 GCACCCAACTGATCTTACGATCTTTTACTTTTCCACGCGTTTCTGGGTGAGCAAAAACA 10138

Db 8999 AACTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACC 9058  
Qy 6793 TTCCGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGT 6852  
Db 9059 TTCCGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGT 9118  
Qy 6853 TTTTGTGTTCCAAAGCAGCAGATTACCGGAGAAAAAGGATCTCAAGAGATCTCTTTG 6912  
Db 9119 TTTTGTGTTCCAAAGCAGCAGATTACCGGAGAAAAAGGATCTCAAGAGATCTCTTTG 9178  
Qy 6913 ATCTTTTCTACGGGCTGTGAGCTCAGTGAAGCAAAACTCACGTTAAGGATTTTGGTC 6972  
Db 9179 ATCTTTTCTACGGGCTGTGAGCTCAGTGAAGCAAAACTCACGTTAAGGATTTTGGTC 9238  
Qy 6973 ATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAATTTAAATTTAA 7032  
Db 9239 ATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAATTTAAATTTAA 9298  
Qy 7033 TCAATCTAAAAGTATATAGTAAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 7092  
Db 9299 TCAATCTAAAAGTATATAGTAAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 9358  
Qy 7093 GCACCTATCTCAGCGATCTGTCTATTTTCGTTTCACTCCATAGTGTGCTGACTCCCGT 7152  
Db 9359 GCACCTATCTCAGCGATCTGTCTATTTTCGTTTCACTCCATAGTGTGCTGACTCCCGT 9418  
Qy 7153 TAGATAAATAAGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 7212  
Db 9419 TAGATAAATAAGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 9478  
Qy 7213 GACCAACGCTCAGCGCTCCAGATTTATCAGCAATATAACCAAGCCAGCGGAGCCGAG 7272  
Db 9479 GACCAACGCTCAGCGCTCCAGATTTATCAGCAATATAACCAAGCCAGCGGAGCCGAG 9538  
Qy 7273 CGCAGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAA 7332  
Db 9539 CGCAGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAA 9598  
Qy 7333 GCTAGATTAAGTGTGCTCCAGTAAATAGTTTGGCAACGTTTGTCCATTTGCTGCAGGC 7392  
Db 9599 GCTAGATTAAGTGTGCTCCAGTAAATAGTTTGGCAACGTTTGTCCATTTGCTGCAGGC 9658  
Qy 7393 ATCTGTGTGTCAAGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGTTCCCAAGATCA 7452  
Db 9659 ATCTGTGTGTCAAGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGTTCCCAAGATCA 9718  
Qy 7453 AGCGAGTTATATGATCCCCATGTTGTGCAAAAAGCGGTAGTCTCTCGTCCCTCCG 7512  
Db 9719 AGCGAGTTATATGATCCCCATGTTGTGCAAAAAGCGGTAGTCTCTCGTCCCTCCG 9778  
Qy 7513 ATCTGTGTGTCAGAAATAGTTGGCCGAGTTATCACTCATGTTTATGGCAGCACTGCAT 7572  
Db 9779 ATCTGTGTGTCAGAAATAGTTGGCCGAGTTATCACTCATGTTTATGGCAGCACTGCAT 9838  
Qy 7573 AATCTCTTATCTGATCCGATCCGATAGATGCTTTCTGTGACTGTGTGACTCAACC 9898  
Db 9839 AATCTCTTATCTGATCCGATCCGATAGATGCTTTCTGTGACTGTGTGACTCAACC 9958  
Qy 7633 AAGTCAATCTGAGATAGTGTATGCGGAGCCAGTGTGCTCTTGGCCGCGCTCAACACGG 7692  
Db 9899 AAGTCAATCTGAGATAGTGTATGCGGAGCCAGTGTGCTCTTGGCCGCGCTCAACACGG 9958  
Qy 7693 GATAATACCGCCCATAGCAGAACTTTAAAGTGTCTCATCTTTGGAAGAAAGTCTTTCG 7752  
Db 9959 GATAATACCGCCCATAGCAGAACTTTAAAGTGTCTCATCTTTGGAAGAAAGTCTTTCG 10018  
Qy 7753 GGGCGAAATCTCTAAGGATCTTACCGTGTGTGAGATCCAGTTCAGTGTAAACCACTCGT 7812  
Db 10019 GGGCGAAATCTCTAAGGATCTTACCGTGTGTGAGATCCAGTTCAGTGTAAACCACTCGT 10078  
Qy 7813 GCACCCAACTGATCTTACGATCTTTTACTTTTCCACGCGTTTCTGGGTGAGCAAAAACA 7872  
Db 10079 GCACCCAACTGATCTTACGATCTTTTACTTTTCCACGCGTTTCTGGGTGAGCAAAAACA 10138

QY 7873 GGAAGGCAAAATGCCGCAAAAAGGGAATAAGGCGGACACGAAATGTTGAATCTCAT 7932  
 Db 10139 GGAAGGCAAAATGCCGCAAAAAGGGAATAAGGCGGACACGAAATGTTGAATCTCAT 10198  
 QY 7933 CTCCTCTCTTTTCAATATATTGAGCATTTATCAGGGTTATCTCATGAGCGGATAC 7992  
 Db 10199 CTCCTCTCTTTTCAATATATTGAGCATTTATCAGGGTTATCTCATGAGCGGATAC 10258  
 QY 7993 ATATTGAAATGATTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCGCAAAA 8052  
 Db 10259 ATATTGAAATGATTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCGCAAAA 10318  
 QY 8053 GTGCCACCTGAGGCTTAAGAACCAATATTATCATGCAATTAACCTATAAATAAGGGGT 8112  
 Db 10319 GTGCCACCTGAGGCTTAAGAACCAATATTATCATGCAATTAACCTATAAATAAGGGGT 10378  
 QY 8113 ATCAGAGGCGCTTTGCTCTTCAAGAAATTAATCTCATGTTGACAGCTTATCATCGATA 8172  
 Db 10379 ATCAGAGGCGCTTTGCTCTTCAAGAAATTAATCTCATGTTGACAGCTTATCATCGATA 10438  
 QY 8173 AGCTGACTCATGTTGGTATTGTAATAGAGCGAGATCGGAAACACTGAAATAACAGT 8232  
 Db 10439 AGCTGACTCATGTTGGTATTGTAATAGAGCGAGATCGGAAACACTGAAATAACAGT 10498  
 QY 8233 TATTATTGG 8241  
 Db 10499 TATTATTGG 10507

RESULT 12  
 US-10-038-722-71  
 ; Sequence 71, Application US/10038722  
 ; Publication No. US20030175919A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LEY, Arthur C.  
 ; APPLICANT: GUTERMAN, Sonia K.  
 ; APPLICANT: MARKLAND, William  
 ; APPLICANT: KENT, Rachel B.  
 ; APPLICANT: ROBERTS, Bruce L.  
 ; APPLICANT: LADNER, Robert K.  
 ; TITLE OF INVENTION: ITI-D1 KUNITZ DOMAIN MUTANTS AS nHE INHIBITORS  
 ; FILE REFERENCE: LEY-1B  
 ; CURRENT APPLICATION NUMBER: US/10/038,722  
 ; CURRENT FILING DATE: 2002-01-08  
 ; PRIOR APPLICATION NUMBER: US 08/849,406  
 ; PRIOR FILING DATE: 1998-07-21  
 ; PRIOR APPLICATION NUMBER: PCT/US95/16349  
 ; PRIOR FILING DATE: 1995-12-15  
 ; PRIOR APPLICATION NUMBER: US 08/358,160  
 ; PRIOR FILING DATE: 1994-12-16  
 ; NUMBER OF SEQ ID NOS: 129  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 71  
 ; LENGTH: 8584  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Plasmid pHIL-D2 (MfalpaphPrePro::EPI-HNE-3) (table 251)  
 US-10-038-722-71

Query Match 62.4%; Score 5142.8; DB 15; Length 8584;  
 Best Local Similarity 94.6%; Pred. No. 0;  
 Matches 5535; Conservative 0; Mismatches 167; Indels 152; Gaps 14;  
 QY 2 GATCTAACATCCAAAGCAGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 61  
 Db 14 GATCTAACATCCAAAGCAGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 73  
 QY 62 TCATTTCTCACATTAAGTCCAAACGCAACAGGAGGGGATACATAGCAGCAGACCGTT 121  
 Db 74 TCATTTCTCACATTAAGTCCAAACGCAACAGGAGGGGATACATAGCAGCAGACCGTT 133

QY 122 GCAACGCGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGGCCATCGAAAAACA 181  
 Db 134 GCAACGCGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGGCCATCGAAAAACA 193  
 QY 182 GCCCAGCTTATTGGGCTTGATTGAGCTCGCTCATTTCCAAATCTCTTATTAGGCTACTAA 241  
 Db 194 GCCCAGCTTATTGGGCTTGATTGAGCTCGCTCATTTCCAAATCTCTTATTAGGCTACTAA 253  
 QY 242 CACCATGACTTTATTAGCCTGCTATCTCTGCGCCCTCGCGAGGTTCAATGTTGTTTAT 301  
 Db 254 CACCATGACTTTATTAGCCTGCTATCTCTGCGCCCTCGCGAGG-TCAATGTTGTTTAT 312  
 QY 302 TTCGGAATGCAACAAGCTCCGATTACACCCGAAACATCACTCCAGATGAGGCTTTCTGA 361  
 Db 313 TTCGGAATGCAACAAGCTCCGATTACACCCGAAACATCACTCCAGATGAGGCTTTCTGA 372  
 QY 362 GTGTGGGCTCAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAAGCGTG 421  
 Db 373 GTGTGGGCTCAATAGTTTCATGTT-CCCAATGGCCCAAACTGACAGTTTAAAGCGTG 431  
 QY 422 TCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAGATGAACTAAAGTTTGGTTCGT 481  
 Db 432 TCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAGATGAACTAAAGTTTGGTTCGT 491  
 QY 482 TGAATGCTAAACGCGCAGTTGGTCAAAAAGAAATTTCCAAAAGTCCGCCATACCGTTTGTG 541  
 Db 492 TGAATGCTAAACGCGCAGTTGGTCAAAAAGAAATTTCCAAAAGTCCGCCATACCGTTTGTG 551  
 QY 542 TTGTTTGGTATTGATTGACGATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 601  
 Db 552 TTGTTTGGTATTGATTGACGATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 611  
 QY 602 TCTATCGCTTCTGAAACCCCGGTGCACCTGTGCGGAAACGCAAAATGGGGAAC-ACCCTGT 660  
 Db 612 TCTATCGCTTCTGAAACCCCGGTGCACCTGTGCGGAAACGCAAAATGGGGAACCAACCCGT 671  
 QY 661 TTTTGGATGATTATGATTTGT-CTCACATTTGATGCTTCCAGATCTGTTGGGATAC 719  
 Db 672 TTTTGGATGATTATGATTTGTCTCCACATTTGATGCTTCCAGATCTGTTGGGATAC 731  
 QY 720 TGCTGATAGCTAAACGTTTCAATGATCAAAATTTAACTGTTCTCAACCTACTTGACA-GCA 778  
 Db 732 TGCTGATAGCTAAACGTTTCAATGATCAAAATTTAACTGTTCTCAACCTACTTGACAGGCA 791  
 QY 779 ATATATAACAGAGGAGCTGCCCTGTCTTAAACCTTTTATTTTATCATATTATAGC 838  
 Db 792 ATATATAACAGAGGAGCTGCCCTGTCTTAAACCTTTTATTTTATCATATTATAGC 851  
 QY 839 TTACTTTTCAATTTGGGCTGTTCCAAATGCAAGCTTTTGAATTTTAAAGCTTTTAAAC 898  
 Db 852 TTACTTTTCAATTTGGGCTGTTCCAAATGCAAGCTTTTGAATTTTAAAGCTTTTAAAC 911  
 QY 899 GACAACTTGAGAGATCAAAAAACAACCTAATTTTGAAGGATCCAAACGATCAGATTTTC 958  
 Db 912 GACAACTTGAGAGATCAAAAAACAACCTAATTTTGAAGGATCCAAACGATCAGATTTTC 963  
 QY 959 CTTCAATTTTACTGCAAGTTTATTGCGAGCATCTCCGATTTAGCTGCTCCAGTCAACA 1018  
 Db 964 CATCTATCTTCACTGCTGTTTGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1023  
 QY 1019 CTACAACAGAGATGAAACGCGCAAAATTCGGCTGAAGCTGTCATCGGTTACTTCAGATT 1078  
 Db 1024 CCCTACTGAGAGCAGACTGCTCAAAATCTCTGCTGAGGCTGATCGGTTACTCTGACT 1083  
 QY 1079 TAGAAGGGATTTGATGTTGCTGTTTGGCCATTTTCCAAACAGCACAATAAAGCGGTTAT 1138  
 Db 1084 TGAAGGCTGATCTCGAGCTGCTGTTTGGCCATTTCTTAACCTCTACTTAAACAACCGTTTGT 1143  
 QY 1139 TGTTTTAAATATCTACTATTGCGAGCATTTCTGTAAGAGAGGAGGTTATCTCTCGAGA 1198  
 Db 1144 TGTTCATCAACACTTACCATCGCTTCTATCTGCTGTAAGAGGAGGTTTCTCTTGACA 1203  
 QY 1199 AAGAGAGGCTGAAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGGAAACAAATGATACA 1258

Db 1204 AGAGAGCTGCTGTAAAC----- 1220  
Qy 1259 TGGCCCCCGGAGAGGATAAATTGTGGCTTCCCGGTGTCCACGCCGACAGTGCACGG 1318  
Db 1221 TTGCCAATCGTCAGAGGTCATGCAATGCTTCTTCCCAAGATGGGCTTTCGAGCC--- 1276  
Qy 1319 AGAGAGGTGTGTTTGTATGACAGTGTCCGGGATCCCGTGTGCTTCCACCCCATGG 1378  
Db 1277 -----TGTTAAGGGTAAGTGCCTGTGTTCCCATACGGTGG 1312  
Qy 1379 CCATCGAGAACACCTCAAGAGAGAAATGTCCTTCTAACTAGTGGCGGTAGAATTCCTCTAG 1438  
Db 1313 TTGTCAAGGTAAC---GGTAACAAGTTCTACTCTGAGAAGAGGTGTAGAGAGTACTGTGG 1369  
Qy 1439 GCGGCCCGGAAATTAATTGCGCTTAGACATGACTGTTCCTCAGTTCAGTTCGACGCTTA 1498  
Db 1370 TGTTCCTATAGTAAGAAATTCGCTTAGACATGACTGTTCCTCAGTTCAGTTCGAGCA-TTA 1428  
Qy 1499 CGAGAGAGCCGCTTGTGATGATCTTAATCAAGAGGATGTGAGAAATGCCAATTTGCCCTGAG 1558  
Db 1429 CGAGAGAGCCGCTTGTGATGATCTTAATCAAGAGGATGTGAGAAATGCCAATTTGCCCTGAG 1488  
Qy 1559 AGATCGAGGCTCAATTTTGTATGATCTTTTATTTTGTAACTATATAGTATAGGATTTTTT 1618  
Db 1489 AGATCGAGGCTCAATTTTGTATGATCTTTTATTTTGTAACTATATAGTATAGGATTTTTT 1548  
Qy 1619 TTGTCATTTTGTCTTCTCGTACGAGCTTCTGATCAGCCTATCTCGCAGCTCATG 1678  
Db 1549 TTGTCATTTTGTCTTCTCGTACGAGCTTCTGATCAGCCTATCTCGCAGCTCATG 1608  
Qy 1679 AATATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGAATTTTCTTGTGATTTTC 1738  
Db 1609 AATATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGAATTTTCTTGTGATTTTC 1668  
Qy 1739 CCACCTCTTTCAGATACAGAGATTAAGTGAGAGTTGCTTGTTCGACGCTTATCGAT 1798  
Db 1669 CCACCTCTTTCAGATACAGAGATTAAGTGAGAGTTGCTTGTTCGACGCTTATCGAT 1728  
Qy 1799 AAGCTTTAATCGGTAGTTTATCAGATTAAATGCTAAACGAGTCAAGGACCGGTGATG 1858  
Db 1729 AAGCTTTAATCGGTAGTTTATCAGATTAAATGCTAAACGAGTCAAGGACCGGTGATG 1788  
Qy 1859 AATCTAACAATGCGCTCATGCTATCTCGGACCGTCAACCTCGGATGCTGAGGATA 1918  
Db 1789 AATCTAACAATGCGCTCATGCTATCTCGGACCGTCAACCTCGGATGCTGAGGATA 1848  
Qy 1919 GGTCTGTGTTATGCGGTACTGCGGGCTCTTGGGGATATCGTCCATTCGACAGCATC 1978  
Db 1849 GGTCTGTGTTATGCGGTACTGCGGGCTCTTGGGGATATCGTCCATTCGACAGCATC 1908  
Qy 1979 GCCACTCATATGCGGTCTCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCC 2038  
Db 1909 GCCACTCATATGCGGTCTCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCC 1968  
Qy 2039 GTTCTCGAGACATGTCGCGCGCTTTGGCGCGCGCGCGCGCTCTGCTGCTTCTGCTACTT 2098  
Db 1969 GTTCTCGAGACATGTCGCGCGCTTTGGCGCGCGCGCGCGCTCTGCTGCTTCTGCTACTT 2028  
Qy 2099 GGAGCCACTATCGACTACGCGATCATGGCGACCAACCCGCTCTGTGTGATCTATPCGAATC 2158  
Db 2029 GGAGCCACTATCGACTACGCGATCATGGCGACCAACCCGCTCTGTGTGATCTATPCGAATC 2088  
Qy 2159 TAAATGTAAGTTAAATCTCTAAATAATTAATTAAGTCCGAGTTTCTCCATACGAACCTT 2218  
Db 2089 TAAATGTAAGTTAAATCTCTAAATAATTAATTAAGTCCGAGTTTCTCCATACGAACCTT 2148  
Qy 2219 AACACATTTGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCCATCTGCTTGGCC 2278  
Db 2149 AACACATTTGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCCATCTGCTTGGCC 2208  
Qy 2279 AATATGTTTCACTCCCTCAGAGTTACGTCTGTGTAAGTGAATGAACTTCTGGAAGGTTGC 2338

Db 2209 AATATGTTTCACTCCCTCAGGAGTTACGTCTTGTGAAGTGAATCTTCTGGAAGGTTGC 2268  
Qy 2339 AGTGTTAACCTCGCTGTATTGACGGGCATATCCGTACGTTGCGAAAGTGTGTTGTTGATCC 2398  
Db 2269 AGTGTTAACCTCGCTGTATTGACGGGCATATCCGTACGTTGCGAAAGTGTGTTGTTGATCC 2328  
Qy 2399 GGAGAGGTAATCTCCACAACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGCGTG 2458  
Db 2329 GGAGAGGTAATCTCCACAACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGCGTG 2388  
Qy 2459 TTGTACTTGTATCAACATTAAGAAGAAGCATTTCCGATTTTGCAGGATCAAGTGTTCAGGAGC 2518  
Db 2389 TTGTACTTGTATCAACATTAAGAAGAAGCATTTCCGATTTTGCAGGATCAAGTGTTCAGGAGC 2448  
Qy 2519 GTACTGATTGGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCCGATAGGGTTGTAGAGTG 2578  
Db 2449 GTACTGATTGGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCCGATAGGGTTGTAGAGTG 2508  
Qy 2579 TCGAATACACTTGTGCTACAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGC 2638  
Db 2509 TCGAATACACTTGTGCTACAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGC 2568  
Qy 2639 ATCTTCAATTTCTGCAAGCTTCTGCTGTCTGATATCGACAGCCACAGAAATCACTCGGGA 2698  
Db 2569 ATCTTCAATTTCTGCAAGCTTCTGCTGTCTGATATCGACAGCCACAGAAATCACTCGGGA 2628  
Qy 2699 ATCAATACACTTGTGCTACAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGC 2758  
Db 2629 ATCAATACACTTGTGCTACAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGC 2687  
Qy 2759 TTTATCAGCAATTAATAGAACTTCAAGAGGCGCCAGCAGCATGTCAATCTACACAGGC 2818  
Db 2688 TTTATCAGCAATTAATAGAACTTCAAGAGGCGCCAGCAGCATGTCAATCTACACAGGC 2747  
Qy 2819 TGAATGTCTCATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTGACCAAAA 2878  
Db 2748 TGAATGTCTCATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTGACCAAAA 2807  
Qy 2879 TATTTTGTACACTTAGGAAACGTTTCTGCTGTAAGCCATAGCAGTACTGCTGGGC 2938  
Db 2808 TATTTTGTACACTTAGGAAACGTTTCTGCTGTAAGCCATAGCAGTACTGCTGGGC 2867  
Qy 2939 GCTCTCTGTACAGCATACACTTAGCACCAACCTTGTGGCAACGATAGATGACTTCTGG 2998  
Db 2868 GCTCTCTGTACAGCATACACTTAGCACCAACCTTGTGGCAACGATAGATGACTTCTGG 2927  
Qy 2999 GGTAAAGGTTACCATCTTCTAGTGGAGATGCAAAAAACAAATTTCTTTGCAACCAAGCAAC 3058  
Db 2928 GGTAAAGGTTACCATCTTCTAGTGGAGATGCAAAAAACAAATTTCTTTGCAACCAAGCAAC 2987  
Qy 3059 TTTGCGAGAACACCCACGATCAGGAGTAGTGGAGGAGAAATGCGGTTCCACAGGAAAT 3118  
Db 2988 TTTGCGAGAACACCCACGATCAGGAGTAGTGGAGGAGAAATGCGGTTCCACAGGAAAT 3047  
Qy 3119 ATAGAGGCGCAACTTCTCAATAGTCTTTCGCAAAACGAGAGCAGACTACACAGGCAAGT 3178  
Db 3048 ATAGAGGCGCAACTTCTCAATAGTCTTTCGCAAAACGAGAGCAGACTACACAGGCAAGT 3107  
Qy 3179 CTTCACTTGCAGAGTCTCGTTAGTTGAGCTTCATGGAATTTCTGAGGTTATCTATAGA 3238  
Db 3108 CTTCACTTGCAGAGTCTCGTTAGTTGAGCTTCATGGAATTTCTGAGGTTATCTATAGA 3167  
Qy 3239 GAGATCAATGGCTCTCTTAAACGTTATCTGGCAATTTGCAATAAGTTCCTCTGGGAAGGAGC 3298  
Db 3168 GAGATCAATGGCTCTCTTAAACGTTATCTGGCAATTTGCAATAAGTTCCTCTGGGAAGGAGC 3227  
Qy 3299 TTTCTAACACAGGTGTCTTCAAGCGACTCCCATCAAACTTTGGCAGTGTAGTCTTAAAGGCG 3358  
Db 3228 TTTCTAACACAGGTGTCTTCAAGCGACTCCCATCAAACTTTGGCAGTGTAGTCTTAAAGGCG 3287  
Qy 3359 TTTGTCCACCATTTTGCAGAACATTTCTGCACAATTTGTTGACTAATTCATTAATCTGTTTC 3418  
Db 3288 TTTGTCCACCATTTTGCAGAACATTTCTGCACAATTTGTTGACTAATTCATTAATCTGTTTC 3347

QY	3419	CGTTTCTGGATAGACGACGAGGGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAC	3478
Db	3348	CGTTTCTGGATAGACGACGAGGGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAC	3407
QY	3479	GTCAATTTTGACAAATCAATACGACCTTCAGAGGAGCTCTTCTTAGGTTTGGATCTTC	3538
Db	3408	GTCAATTTTGACAAATCAATACGACCTTCAGAGGAGCTCTTCTTAGGTTTGGATCTTC	3467
QY	3539	TTTAGGTTGTTCTTGGTGATTCCTTGGCTTGGCATCTCTTCTTCTTAGTGACCTTTAG	3598
Db	3468	TTTAGGTTGTTCTTGGTGATTCCTTGGCTTGGCATCTCTTCTTCTTAGTGACCTTTAG	3527
QY	3599	GGACTTCATATCCAGGTTTCTCTCCACCTCGTCCAACTGACCGTACTTTGGACACATCT	3658
Db	3528	GGACTTCATATCCAGGTTTCTCTCCACCTCGTCCAACTGACCGTACTTTGGACACATCT	3587
QY	3659	AACTAATGCAAAATAAATAAGTACGACATTCGCCAGGCTATATCTTCTTGGATTTAGC	3718
Db	3588	AACTAATGCAAAATAAATAAGTACGACATTCGCCAGGCTATATCTTCTTGGATTTAGC	3647
QY	3719	TTCTCCAGTTTCATCAGCTTCTCTCCCTAATTTTAGCGTTTCAACAAATCTTCGTCTCAAA	3778
Db	3648	TTCTCCAGTTTCATCAGCTTCTCTCCCTAATTTTAGCGTTTCAACAAATCTTCGTCTCAAA	3707
QY	3779	TAAACGTTTGGTATAGAAACCTTCTGGAGCATTTGCTTTAGCATCCCAAGGTGGCTTC	3838
Db	3708	TAAACGTTTGGTATAGAAACCTTCTGGAGCATTTGCTTTAGCATCCCAAGGTGGCTTC	3766
QY	3839	GATGCTCTAAGACCTTTGATGGCCAAACAGAGAGTGGTTCCTAAGTACAGAGAAC	3898
Db	3767	CATGCTCTAAGACCTTTGATGGCCAAACAGAGAGTGGTTCCTAAGTACAGAGAAC	3826
QY	3899	AACACTGTTTGTCTCAACACAAATTTCAAGCAGTCTCATCACAAATCCAATTCGATACC	3958
Db	3827	AACACTGTTTGTCTCAACACAAATTTCAAGCAGTCTCATCACAAATCCAATTCGATACC	3886
QY	3959	CAGCAACTTTTGGTTGCTCCAGATGTAGCACTTTATACCAAAACCGTACAGAGAGA	4018
Db	3887	CAGCAACTTTTGGTTGCTCCAGATGTAGCACTTTATACCAAAACCGTACAGAGAGA	3946
QY	4019	TTGTAGACTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGAGAGTACAC	4078
Db	3947	TTGTAGACTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGAGAGTACAC	4006
QY	4079	CAGGCCCAACGAGTAATAGAGAGTACGCCCAAGAGTGAATAGACCAATCGGGCG	4138
Db	4007	CAGGCCCAACGAGTAATAGAGAGTACGCCCAAGAGTGAATAGACCAATCGGGCG	4066
QY	4139	GTCAAGTACCAAGACGCCCAACAAATTTCACTGACAGGAATTTTGGACATCTCAGA	4198
Db	4067	GTCAAGTACCAAGACGCCCAACAAATTTCACTGACAGGAATTTTGGACATCTCAGA	4126
QY	4199	AAGTTCTGATTCAGTAGTCAATTTCCGAGCATCAATTAATGGGATTTATACCAAGAGCAAC	4258
Db	4127	AAGTTCTGATTCAGTAGTCAATTTCCGAGCATCAATTAATGGGATTTATACCAAGAGCAAC	4186
QY	4259	AGTGGAGTACATCTACCAATTTTGGCTGCTCAGAAAAGCATAAACAGTTCTACTACC	4318
Db	4187	AGTGGAGTACATCTACCAATTTTGGCTGCTCAGAAAAGCATAAACAGTTCTACTACC	4246
QY	4319	GCCATTAGTGAATTTTCAATTCGCCAGTGGAGAGAAAGGCACAGCGATACTAGC	4378
Db	4247	GCCATTAGTGAATTTTCAATTCGCCAGTGGAGAGAAAGGCACAGCGATACTAGC	4306
QY	4379	ATTAGCGGCAAGGATGCAATTTTATCAACCGAGGTCTTATAGATAAACCTTAGCGCTGG	4438
Db	4307	ATTAGCGGCAAGGATGCAATTTTATCAACCGAGGTCTTATAGATAAACCTTAGCGCTGG	4366
QY	4439	GATCATCTTTGACAACTTTCTGCGAATCTAGTCCAAATCACTTCAATTTGATACC	4498
Db	4367	GATCATCTTTGACAACTTTCTGCGAATCTAGTCCAAATCACTTCAATTTGATACC	4426

QY	4499	ATTATTTGTACAACTTGAGCAAGTTGTGATCAGCTCCTCAAAATTGGTCTCTGTACCGA	4558
Db	4427	ATTATTTGTACAACTTGAGCAAGTTGTGATCAGCTCCTCAAAATTGGTCTCTGTACCGA	4436
QY	4559	TGACTCAACTTGGACATTAACCTTGAAGCTCAGTCGATTCAGTGAACTTGATCAGGTTTG	4618
Db	4437	TGACTCAACTTGGACATTAACCTTGAAGCTCAGTCGATTCAGTGAACTTGATCAGGTTTG	4496
QY	4619	CAGTGTGAGCAGCAGTAGGAAACACGCTTTTCTACCAAACTCAAGGAAATATCAAA	4678
Db	4497	CAGTGTGAGCAGCAGTAGGAAACACGCTTTTCTACCAAACTCAAGGAAATATCAAA	4556
QY	4679	CTCTGCAACACTTGGCTATGCGATAGCAGTAGCAAGGAAATGTCTACTTGAAGTCGACAGTG	4738
Db	4557	CTCTGCAACACTTGGCTATGCGATAGCAGTAGCAAGGAAATGTCTACTTGAAGTCGACAGTG	4616
QY	4739	AGTGTAGTCTTGAAGAAATCTGAAAGCGCTATTTTATTTATCAGTGAATCATCAGCA	4798
Db	4617	AGTGTAGTCTTGAAGAAATCTGAAAGCGCTATTTTATTTATCAGTGAATCATCAGCA	4676
QY	4799	GATCCTCTACGCCGCGACGATCGTGGCCGACCTGCAAGTTCGGCATCACCGGCGCCACAG	4858
Db	4677	GATCCTCTACGCCGCGACGATCGTGGC-----CGGCATCACCGGCGCCACAGG	4724
QY	4859	TGCGTTGCTGGCGCTATATCCGCATCACCGATGGGAGATCGGGCTCGCCACTT	4918
Db	4725	TGCGTTGCTGGCGCTATATCCGCATCACCGATGGGAGATCGGGCTCGCCACTT	4784
QY	4919	CGGCTCATGAGCGCTTGTTCGGCGTGGGTATGTGGCAGGCGCGGCGGAGCT	4978
Db	4785	CGGCTCATGAGCGCTTGTTCGGCGTGGGTATGTGGCAGGCGCGGCGGAGCT	4844
QY	4979	GTTGGGCGCATCTCTTGTGATGACCATTCCTTGGCGGCGCGTGTCTAACCGCTCAA	5038
Db	4845	GTTGGGCGCATCTCTTGTGATGACCATTCCTTGGCGGCGCGTGTCTAACCGCTCAA	4904
QY	5039	CCTACTACTGGCTGCTTCTTAATGAGGAGTGCATTAAGCGAGAGCGTGCAGTATCTAT	5098
Db	4905	CCTACTACTGGCTGCTTCTTAATGAGGAGTGCATTAAGCGAGAGCGTGCAGTATCTAT	4964
QY	5099	GATTGGAAGTATGGAAATGGTGTATCCCGCATTTCTCAGTGTCTTGAAGTCTCTATCAG	5158
Db	4965	GATTGGAAGTATGGAAATGGTGTATCCCGCATTTCTCAGTGTCTTGAAGTCTCTATCAG	5024
QY	5159	ATTATGCCCACTTAAGCAACCGGAGGAGATTTTCATGTAATTTCTCTGACTTTTG	5218
Db	5025	ATTATGCCCACTTAAGCAACCGGAGGAGATTTTCATGTAATTTCTCTGACTTTTG	5084
QY	5219	GTCACTCAGTAGACTCGAACTGTGAGACTATCTCGGTTTATGACAGAGAAATGTCTTCTT	5278
Db	5085	GTCACTCAGTAGACTCGAACTGTGAGACTATCTCGGTTTATGACAGAGAAATGTCTTCTT	5144
QY	5279	GGAGACAGTAAATGAAGTCCCAACATTAAGAAATCTCTTGTATCAGGAACAACTTCTT	5338
Db	5145	GGAGACAGTAAATGAAGTCCCAACATTAAGAAATCTCTTGTATCAGGAACAACTTCTT	5204
QY	5339	GTTTCGAACCTTTTTCGGTGTCTGAACTATAAAATGTAGAGTGGATATGTCCGGTAGGAA	5398
Db	5205	GTTTCGAACCTTTTTCGGTGTCTGAACTATAAAATGTAGAGTGGATATGTCCGGTAGGAA	5264
QY	5399	TGGAGCGGCAAAATGCTTCTTCTGCACTTCAAGAGGTATGTAGGTTTGTAGTACT	5458
Db	5265	TGGAGCGGCAAAATGCTTCTTCTGCACTTCAAGAGGTATGTAGGTTTGTAGTACT	5324
QY	5459	GATGCCAACTTCAGTGCAAGCTGCTATTTTCGTTCAAACTTCGAATCCAGAGAAAT	5518
Db	5325	GATGCCAACTTCAGTGCAAGCTGCTATTTTCGTTCAAACTTCGAATCCAGAGAAAT	5384
QY	5519	CAAAAGTTTGTCTACTATTTGATCCAAAGCAGTGGCTCTTGAATTCGAATAGTGTG	5578
Db	5385	CAAAAGTTTGTCTACTATTTGATCCAAAGCAGTGGCTCTTGAATTCGAATAGTGTG	5444
QY	5579	CTCGTTTGTGAGTCACTTTTGTATGAATAATCTAGTCTTTCATCTAAATATCTTGA	5638





1379 CCATCGAGAACACTCAAGAGAGAAATGTCCCTTCTAACTAGTGGCGTAGAAATTCCTTAG 1438  
1313 TTGTCAGAGGTAAC- - -GGTAACAAGTCTACTCTGAGAAGGAGTGTPAGAGACTACTGTGG 1369  
1439 GCGCGCGCGAATTAATTCGCTCTAGACATGACTGTTCTCTAGTTCAGTTCAGGTCGACTTA 1498  
1370 TGTTCATAGTAAGAAATTCGCTCTAGACATGACTGTTCTCTAGTTCAGTTCAGGTCGACTTA 1428  
1499 CGAAGAACCGGTCCTTCTAGATTTCTAAATCAAGAGGATGTCAAGATGCCATTTGGCTGAG 1558  
1429 CGAAGAACCGGTCCTTCTAGATTTCTAAATCAAGAGGATGTCAAGATGCCATTTGGCTGAG 1488  
1559 AGATGAGGCTTCATTTTGTATACATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1618  
1489 AGATGAGGCTTCATTTTGTATACATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1548  
1619 TTGTCTATTTTGTCTTCTCTGACGAGCTTGCTCTGATCAGCTATCTCGAGCTGATG 1678  
1549 TTGTCTATTTTGTCTTCTCTGACGAGCTTGCTCTGATCAGCTATCTCGAGCTGATG 1608  
1679 AATATCTTGTGTAGGGGTTTGGGAAATCATTTGAGTTTGTAGTTTCTTTGGTATTTTC 1738  
1609 AATATCTTGTGTAGGGGTTTGGGAAATCATTTGAGTTTGTAGTTTCTTTGGTATTTTC 1668  
1739 CCACTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTTCGTTTGTGCAAGCTTATCGAT 1798  
1669 CCACTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTTCGTTTGTGCAAGCTTATCGAT 1728  
1799 AAGCTTTAATGCGGTAGTTTATCACAGTTTAAATTTGCTAACGAGTCAGGACCGGTATG 1858  
1729 AAGCTTTAATGCGGTAGTTTATCACAGTTTAAATTTGCTAACGAGTCAGGACCGGTATG 1788  
1859 AATATCTAAGTCGCTCATCTGTCATCTCGGACCGTCACCTGATGCTGTAGGCATA 1918  
1789 AATATCTAAGTCGCTCATCTGTCATCTCGGACCGTCACCTGATGCTGTAGGCATA 1848  
1919 GCGTTGGTTATGCGGTACTGCGGGCCCTTTCGCGGATATCGTCCATTCGACAGCATC 1978  
1849 GCGTTGGTTATGCGGTACTGCGGGCCCTTTCGCGGATATCGTCCATTCGACAGCATC 1908  
1979 GCGAGTCACTATGCGGTGCTGCTAGGCTATATGCGGTATGATGCAATTTCTATGCGCACCC 2038  
1909 GCGAGTCACTATGCGGTGCTGCTAGGCTATATGCGGTATGATGCAATTTCTATGCGCACCC 1968  
2039 GTTCTCGGAGCACTGTCGACCGCTTTGCGCGCGCCGACGTCCTGCTCGCTTCTGCTACTT 2098  
1969 GTTCTCGGAGCACTGTCGACCGCTTTGCGCGCGCCGACGTCCTGCTCGCTTCTGCTACTT 2028  
2099 GGAGCCACTATCGACTACCGGATCATGCGGACCAACCGGTCCTGCGATCTATCGAATC 2158  
2029 GGAGCCACTATCGACTACCGGATCATGCGGACCAACCGGTCCTGCGATCTATCGAATC 2088  
2159 TAAATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2218  
2089 TAAATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2148  
2219 AACAGCAITTCGCGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCC 2278  
2149 AACAGCAITTCGCGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCC 2208  
2279 AATATGTTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTATGAACTTCTGGAAGTTTC 2338  
2209 AATATGTTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTATGAACTTCTGGAAGTTTC 2268  
2339 AGTGTAACTCCGCTGTATTTAGCGGCATATCCGTACGTTGGCAAGTGTGTTGGTACC 2398  
2269 AGTGTAACTCCGCTGTATTTAGCGGCATATCCGTACGTTGGCAAGTGTGTTGGTACC 2328  
2399 GGAGAGTAACTCCACAACTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTG 2458  
2329 GGAGAGTAACTCCACAACTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTG 2388

2459 TTGTACTGTATCAACATAAGAAGAGCATTTCTCGATTTGCAGGATCAAGTGTTCAGGAGC 2518  
2389 TTGTACTGTATCAACATAAGAAGAGCATTTCTCGATTTGCAGGATCAAGTGTTCAGGAGC 2448  
2519 GTACTGTATGACATTTTCCAAAGCCTGCTGCTAGTTCGACCGATAGGTTGTAGAGTG 2578  
2449 GTACTGTATGACATTTTCCAAAGCCTGCTGCTAGTTCGACCGATAGGTTGTAGAGTG 2508  
2579 TSCAATACACTTGGTGTCAATTTTCAACCCCTTGGCAACTGTCACAGCTTGGTTGTGAACAGC 2638  
2509 TSCAATACACTTGGTGTCAATTTTCAACCCCTTGGCAACTGTCACAGCTTGGTTGTGAACAGC 2568  
2639 ATCTTCAATTTCTGCAAGCTTCTTGTCTGTATATGACAGCCAAACAGATCACTCGGA 2698  
2569 ATCTTCAATTTCTGCAAGCTTCTTGTCTGTATATGACAGCCAAACAGATCACTCGGA 2628  
2699 ATCAATACACTTGGTGTGAGACAGAGCTGTGAGGCAACGAAATCTGATCAGCGTA 2758  
2629 ATCAATACACTTGGTGTGAGACAGAGCTGTGAGGCAACGAAATCTGATCAGCGTA 2687  
2759 TTATATCAAGCAATATCTAGAACTTTCAGAAAGCCCAAGCAGGATGTCAATATACACAGGC 2818  
2688 TTATATCAAGCAATATCTAGAACTTTCAGAAAGCCCAAGCAGGATGTCAATATACACAGGC 2747  
2819 TGATGTGTCAATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGTTTCTCGACCAAA 2878  
2748 TGATGTGTCAATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGTTTCTCGACCAAA 2807  
2879 TATTTTGTCACTTTAGAAACAGTTTCTGTTCCGTAAAGCCATAGCAGCTACTGCTGGGC 2938  
2808 TATTTTGTCACTTTAGAAACAGTTTCTGTTCCGTAAAGCCATAGCAGCTACTGCTGGGC 2867  
2938 GCTTCTCTGCTAGCAGTATACATTTAGCAACCACTTGTGGGCAACGTTAGTACTTCTGG 2998  
2868 GCTTCTCTGCTAGCAGTATACATTTAGCAACCACTTGTGGGCAACGTTAGTACTTCTGG 2927  
2999 GGTAAAGGTTACCATCTTCTTAGTGAGATGCAAAAACAAATTTCTTTGCAACAGCAAC 3058  
2928 GGTAAAGGTTACCATCTTCTTAGTGAGATGCAAAAACAAATTTCTTTGCAACAGCAAC 2987  
3059 TTTGGCAGGAAACCCAGCATCAGGAAAGTGGAAAGCAGAAATTTGCGGTTCCACAGGAAT 3118  
2988 TTTGGCAGGAAACCCAGCATCAGGAAAGTGGAAAGCAGAAATTTGCGGTTCCACAGGAAT 3047  
3119 ATAGAGCCAACTTTCTCAATAGTCTTTCAAACAGAGAGCAGATACACAGGCAAGT 3178  
3048 ATAGAGCCAACTTTCTCAATAGTCTTTCAAACAGAGAGCAGATACACAGGCAAGT 3107  
3179 CTCAACTTGAACCGTCTTCTAGTTAGCTTTCATGGAATTTCTGACGTTATCTATAGA 3238  
3108 CTCAACTTGAACCGTCTTCTAGTTAGCTTTCATGGAATTTCTGACGTTATCTATAGA 3167  
3239 GAGATCAATGGCTCTCTTAACGTTATCTGCAATTTGCAATTTGCTTGGGAAAGAGC 3298  
3168 GAGATCAATGGCTCTCTTAACGTTATCTGCAATTTGCTTGGGAAAGAGC 3227  
3299 TTCTAAACAGAGTGTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAAGGCG 3358  
3228 TTCTAAACAGAGTGTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAAGGCG 3287  
3359 TTTGTACCAATTTTGAAGAACTTTGCGCAATTTGTTTGAATTAATTCATATCTGTTTC 3418  
3288 TTTGTACCAATTTTGAAGAACTTTGCGCAATTTGTTTGAATTAATTCATATCTGTTTC 3347  
3419 CGTTTCTGATAGGAGCAGGAGGCTTCTCAATTTCTTGTGAGAGGCTTTAGAAAC 3478  
3348 CGTTTCTGATAGGAGCAGGAGGCTTCTCAATTTCTTGTGAGAGGCTTTAGAAAC 3407  
3479 GTCAATTTTGCACAAATTCATACGACTTCAGAAAGGAGCTTTCTTTAGGTTTGAATCTTC 3538  
3408 GTCAATTTTGCACAAATTCATACGACTTCAGAAAGGAGCTTTCTTTAGGTTTGAATCTTC 3467  
3539 TTTAGGTTTGTCTTGGTGTATCTCTGGCTTGGCAATCTCTTTTCTTCTTAGTGACCTTTAG 3598



Qy	5757	GCATCTATCTTCTTTTAGAGAAATTTGGGAGATCGGATATCGAGAAAAAGGTAGCTGAT	5816
Db	5625	GCATCTATCTTCTTTTAGAGAAATTTGGGAGATCGGATATCGAGAAAAAGGTAGCTGAT	5684
Qy	5817	TTTAAACGTGAAATTTATCTCAAGATCTCTGCCTCG	5852
Db	5685	TTTAAACGTGAAATTTATCTCAAGATCGGCGCGG	5720
RESULT 14			
US-10-038-722-70			
; Sequence 70, Application US/10038722			
; Publication No. US20030175919A1			
; GENERAL INFORMATION:			
; APPLICANT: LEY, Arthur C.			
; APPLICANT: GUTERMAN, Sonia K.			
; APPLICANT: MARKLAND, William			
; APPLICANT: KENT, Rachel B.			
; APPLICANT: ROBERTS, Bruce L.			
; APPLICANT: LADNER, Robert C.			
; TITLE OF INVENTION: ITI-D1 KUNITZ DOMAIN MUTANTS AS pHE INHIBITORS			
; FILE REFERENCE: LEY-1B			
; CURRENT APPLICATION NUMBER: US/10/038,722			
; CURRENT FILING DATE: 2002-01-08			
; PRIOR APPLICATION NUMBER: US 08/849,406			
; PRIOR FILING DATE: 1999-07-21			
; PRIOR APPLICATION NUMBER: PCT/US95/16349			
; PRIOR FILING DATE: 1995-12-15			
; PRIOR APPLICATION NUMBER: US 08/358,160			
; PRIOR FILING DATE: 1994-12-16			
; NUMBER OF SEQ ID NOS: 129			
; SOFTWARE: Patent in version 3.1			
; SEQ ID NO 70			
; LENGTH: 8157			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Plasmid pHIL-D2 (Table 250)			
US-10-038-722-70			
Query Match 55.4%; Score 4565.2; DB 15; Length 8157;			
Best Local Similarity 89.9%; Pred. No. 0;			
Matches 5262; Conservative 0; Mismatches 13; Indels 579; Gaps 11;			
Qy	2	GATCTAATATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAGG	61
Db	14	GATCTAATATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAGG	73
Qy	62	TCCATTTCTACACATAGTGCCAAACCCACACAGGAGGGATACACTAGCAGCAGACGGT	121
Db	74	TCCATTTCTACACATAGTGCCAAACCCACACAGGAGGGATACACTAGCAGCAGACGGT	133
Qy	122	GCAAAACGAGGACCTCCACACTCTCTCTCTCTCAACACCCACTTTTGGCATCGAAAAACCA	181
Db	134	GCAAAACGAGGACCTCCACACTCTCTCTCTCTCAACACCCACTTTTGGCATCGAAAAACCA	193
Qy	182	GCCAGATTATGGCTTGANTGGAGCTCGCTCATTTCCAAATCTCTTATTAGGCTACTAA	241
Db	194	GCCAGATTATGGCTTGANTGGAGCTCGCTCATTTCCAAATCTCTTATTAGGCTACTAA	253
Qy	242	CACCATGACTTTATTAGCCTGTCTATCCTCGCCCCCTCGCGAGGTTTCATGTTGTTTAT	301
Db	254	CACCATGACTTTATTAGCCTGTCTATCCTCGCCCCCTCGCGAGG-TCATGTTGTTTAT	312
Qy	302	TTCCGAATCAACAAGCTCCGCATTAACCCGCAACATCACTCCAGATGAGGGCTTTCTGA	361
Db	313	TTCCGAATCAACAAGCTCCGCATTAACCCGCAACATCACTCCAGATGAGGGCTTTCTGA	372
Qy	362	GTGTGGGTCAAATAGTTTTCATGTTCCCAAAATGGCCCCCAAACTGACAGTTTAAACGCTG	421
Db	373	GTGTGGGTCAAATAGTTTTCATGTTTCCCAAAATGGCCCCCAAACTGACAGTTTAAACGCTG	431

1002	CGAAGACCGGCTTGTGTAAGTTCTAATCAAGAGATGTCAGATGCCATTGCGCTGAG	1061
1559	AGATGAGCGCTTCATTTTTTGATACATTTTATTTGTAACTATATAGTATAGATTTTTT	1618
1062	AGATGAGCGCTTCAATTTTGATACATTTTATTTGTAACTATATAGTATAGATTTTTT	1121
1619	TTGTCAATTTGTTCTTCTCTGACAGAGTTGCTCTGTATCAGCCTATCTCGAGCTGATG	1678
1122	TTGTCAATTTGTTCTTCTCTGACAGAGTTGCTCTGTATCAGCCTATCTCGAGCTGATG	1181
1679	AATATCTTGTAGGGTTTTGGAAATCATTGAGTTTGATGTTTTCTTGGTATTTC	1738
1182	AATATCTTGTAGGGTTTTGGAAATCATTGAGTTTGATGTTTTCTTGGTATTTC	1241
1739	CGACTCTCTTCAGAGTACAGAAGATTAAAGTGAGAGTTGGTTGTGCAAGCTTATCGAT	1798
1242	CGACTCTCTTCAGAGTACAGAAGATTAAAGTGAGAGTTGGTTGTGCAAGCTTATCGAT	1301
1799	AAGCTTTAATCGGTAGTTTATCAGCTTAAATTTGTAAACGAGTCAGCACCGTGTATG	1858
1302	AAGCTTTAATCGGTAGTTTATCAGCTTAAATTTGTAAACGAGTCAGCACCGTGTATG	1361
1859	AAATCTTAACAATCGCTCATCGTCCATCGGACCGCTCACCCTGGATGCTGTAGGANTA	1918
1362	AAATCTTAACAATCGCTCATCGTCCATCGGACCGCTCACCCTGGATGCTGTAGGANTA	1421
1919	GGCTGGTTATGCGGTACTCGGGGCTCTTCGGGATATCGTCCAATCCGACAGCATC	1978
1422	GGCTGGTTATGCGGTACTCGGGGCTCTTCGGGATATCGTCCAATCCGACAGCATC	1481
1979	GCAGTCACTATGGGTGCTCTAGCGCTATATGCGTTGATCAATTTCTATGCGCACCC	2038
1482	GCAGTCACTATGGGTGCTCTAGCGCTATATGCGTTGATCAATTTCTATGCGCACCC	1541
2039	GTTCTCGGACATGTTCGACCGCTTTGGCGCGGCCAGTCTCTCGCTTCGCTACTT	2098
1542	GTTCTCGGACATGTTCGACCGCTTTGGCGCGGCCAGTCTCTCGCTTCGCTACTT	1601
2099	GGAGCCACTATCGACTACGGATCATGCGGACCAACCGCTCTGTGATCTATCGAATC	2158
1602	GGAGCCACTATCGACTACGGATCATGCGGACCAACCGCTCTGTGATCTATCGAATC	1661
2159	TAAATGTTAAGTTTAAATCTCTAAATTAATTAAGTCCAGTCTCTCCATFAGAACTT	2218
1662	TAAATGTTAAGTTTAAATCTCTAAATTAATTAAGTCCAGTCTCTCCATFAGAACTT	1721
2219	AACAGCATTCGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCATCACTGCTGGCC	2278
1722	AACAGCATTCGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCATCACTGCTGGCC	1781
2279	AATATGTTTCAGTCCCTCAGAGTTAGTCTTTGTGAAGTGATGAACCTTCTGGAAGTTGC	2338
1782	AATATGTTTCAGTCCCTCAGAGTTAGTCTTTGTGAAGTGATGAACCTTCTGGAAGTTGC	1841
2339	AGTGTAACTCCGCTGTATTGAAGGGCATATCGTACGTTTGGCAAAGTGTGTTGATACC	2398
1842	AGTGTAACTCCGCTGTATTGAAGGGCATATCGTACGTTTGGCAAAGTGTGTTGATACC	1901
2399	GGAGAGTAATCTCCACAATCTCTCGAGAGTAGGCCAACAAACAACAGATCCAGCGTG	2458
1902	GGAGAGTAATCTCCACAATCTCTCGAGAGTAGGCCAACAAACAACAGATCCAGCGTG	1961
2459	TTGTACTTTGATCAACATAAGAAAGCATTTCTCGATTTGCAAGATCAAGTGTTCAGAGC	2518
1962	TTGTACTTTGATCAACATAAGAAAGCATTTCTCGATTTGCAAGATCAAGTGTTCAGAGC	2021
2519	GTACTGATTGACATTTCCAAAGCTGCTGTTAGTGTGCAACCGATAGGCTTTAGAGTG	2578
2022	GTACTGATTGACATTTCCAAAGCTGCTGTTAGTGTGCAACCGATAGGCTTTAGAGTG	2081
2579	TGCAATACACTTGGTGACAAATTCACCCCTTGGCAACTGCACAGCTGGTTGTGGAACAGC	2638

Db	2082	TGCAATACACTGGTGAATTTCAACCCCTTGCGAACTGCACAGCTTGGTTGTGAACAGC	2141
Qy	2639	ATCTTCAATTCTGCGAAGCTCCTTGTCTGTATATCGACAGCCAAAGAACTACCTTGGGA	2698
Db	2142	ATCTTCAATTCTGCGAAGCTCCTTGTCTGTATATCGACAGCCAAAGAACTACCTTGGGA	2201
Qy	2699	ATCAATACCATTGTCAGCTTGACAGAGAAGTCTGAGGCAACGAATCTGGATCAGCGTA	2758
Db	2202	ATCAATACCATTGTCAGCTTGAG-CAGAAGGTCTGAGGCAACGAATCTGGATCAGCGTA	2260
Qy	2759	TTTATCAGCAATAACTAGAATCTTCAAGAAGCCCGCAGCAGCATCTCAATACTACACAGGCG	2818
Db	2261	TTTATCAGCAATAACTAGAATCTTCAAGAAGCCCGCAGCAGCATCTCAATACTACACAGGCG	2320
Qy	2819	TGATGTCTCATTTTGACCAATCATCTTGGCAGCAGTAAAGAACTGGTTCCTGGACCAAA	2878
Db	2321	TGATGTCTCATTTTGAAACCAATCATCTTGGCAGCAGTAAAGAACTGGTTCCTGGACCAAA	2380
Qy	2879	TATTTTGTCACTTAGGAACAGTTTCTGTTCGTTAAGCCATAGCAGTACTGCTCTGGCG	2938
Db	2381	TATTTTGTCACTTAGGAACAGTTTCTGTTCGTTAAGCCATAGCAGTACTGCTCTGGCG	2440
Qy	2939	GCCTCTGTAGCAGCATACACTTAGCACCAACCTTGTGGGCAACGATAGATGACTTCTGG	2998
Db	2441	GCCTCTGTGTAGCAGCATACACTTAGCACCAACCTTGTGGGCAACGATAGATGACTTCTGG	2500
Qy	2999	GGTAAGGGTACCATCTCTTTAGTGGAGATGCAAAAACAATTTCTTCCAAACCGACAAC	3058
Db	2501	GGTAAGGGTACCATCTCTTTAGTGGAGATGCAAAAACAATTTCTTCCAAACCGACAAC	2560
Qy	3059	TTTGGCAGGAACACCCAGCATCAGGAAGTGGGAAGCAGAAATTGGGTTCCACACAGGAAT	3118
Db	2561	TTTGGCAGGAACACCCAGCATCAGGAAGTGGGAAGCAGAAATTGGGTTCCACACAGGAAT	2620
Qy	3119	ATAGAGCCCAACTTTCTCAATAGGTCTTGCAAAAACGAGAGCAGCTACACGAGGCAAGT	3178
Db	2621	ATAGAGCCCAACTTTCTCAATAGGTCTTGCAAAAACGAGAGCAGCTACACGAGGCAAGT	2680
Qy	3179	CTCAACTTGCACACGTCTCCGTTAGTTTGGCTTCATCGGAATTTCCCTGACGTTATCTATAGA	3238
Db	2681	CTCAACTTGCACACGTCTCCGTTAGTTTGGCTTCATCGGAATTTCCCTGACGTTATCTATAGA	2740
Qy	3239	GAGATCAATGGCTCTCTTAAAGTTATCTGGCAATTGCATAAGTTCTCTGGGAAAGGAGC	3298
Db	2741	GAGATCAATGGCTCTCTTAAAGTTATCTGGCAATTGCATAAGTTCTCTGGGAAAGGAGC	2800
Qy	3299	TTCTAAACACAGGTGTCTTCAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGCG	3358
Db	2801	TTCTAAACACAGGTGTCTTCAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGCG	2860
Qy	3359	TTTGTCCACCAATTTGAGCAACATTTGCGCAATTTGGTTTGACTAATTCATTAATCTGTTC	3418
Db	2861	TTTGTCCACCAATTTGAGCAACATTTGCGCAATTTGGTTTGACTAATTCATTAATCTGTTC	2920
Qy	3419	CGTTTTCTGATAGGACGACGAAGGCGATCTTCAATTTCTTGTGAGGAGCCCTTAGAAAC	3478
Db	2921	CGTTTTCTGATAGGACGACGAAGGCGATCTTCAATTTCTTGTGAGGAGCCCTTAGAAAC	2980
Qy	3479	GTCAATTTTGCACAATTCATAGGACCTTCAGAAGGCACTTCTTTAGGTTTGGATTCTTTC	3538
Db	2981	GTCAATTTTGCACAATTCATAGGACCTTCAGAAGGCACTTCTTTAGGTTTGGATTCTTTC	3040
Qy	3539	TTTAGGTTGTCTCGGTGTATCTCTGGCTTGGCATCTCCTTTTCCCTTCTAGTGACCTTTAG	3598
Db	3041	TTTAGGTTGTCTCGGTGTATCTCTGGCTTGGCATCTCCTTTTCCCTTCTAGTGACCTTTAG	3100
Qy	3599	GGACTTCATACAGGTTTCTCTCCACCTCGCTCGAAGCTCACACGTACTTTGGCACATCT	3658
Db	3101	GGACTTCATACAGGTTTCTCTCCACCTCGCTCGAAGCTCACACGTACTTTGGCACATCT	3160
Qy	3659	AACTAATGCAAAATAAATAAGTCAGCAATTTCCAGGCTATATCTTCTTGGATTTAGC	3718
Db	3161	AACTAATGCAAAATAAATAAGTCAGCAATTTCCAGGCTATATCTTCTTGGATTTAGC	3220

QY 3719 TTCTGCAAGTTTCATCAGCTTCCTCCCTAATTTTAGCGTTTCAACAAAACCTTCGTCTGCAAA 3778  
DB 3221 TTCTGCAAGTTTCATCAGCTTCCTCCCTAATTTTAGCGTTTCAACAAAACCTTCGTCTGCAAA 3280  
QY 3779 TAACCGTTTGGTATAGAACCTTCCTGGAGCATTCCTTACGATCCCAAGGTGGCTTC 3838  
DB 3281 TAACCGTTTGGTATAGAACCTTCCTGGAGCATTCCTTACGATCCCAAGGT-GCTTC 3339  
QY 3839 CATGGCTTAAGACCTTTGATTGGCCAAAACAGGAAGTGCCTTCCAAAGTGACAGAAACC 3898  
DB 3340 CATGGCTTAAGACCTTTGATTGGCCAAAACAGGAAGTGCCTTCCAAAGTGACAGAAACC 3399  
QY 3899 AACACCTGTTTGTCAACACAAAATTTCAAGCAGTCTCCATCACAAATCCAATTCGATACC 3958  
DB 3400 AACACCTGTTTGTCAACACAAAATTTCAAGCAGTCTCCATCACAAATCCAATTCGATACC 3459  
QY 3959 CAGCAACTTTTCAAGTGTCTCCAGATGTAGCACCTTTTATACCAACAAACCGTGACGACGAGA 4018  
DB 3460 CAGCAACTTTTCAAGTGTCTCCAGATGTAGCACCTTTTATACCAACAAACCGTGACGACGAGA 3519  
QY 4019 TTGGTAGATCCAGTTTGTCTTATAGCTCCCGGAATPAGACTTTTGGACGAGTACAC 4078  
DB 3520 TTGGTAGATCCAGTTTGTCTTATAGCTCCCGGAATPAGACTTTTGGACGAGTACAC 3579  
QY 4079 CAGGCCCAACGAGTAATTAGAGAGTCCAGCACCAAGCTAGTGAATAGACCATCGGGCG 4138  
DB 3580 CAGGCCCAACGAGTAATTAGAGAGTCCAGCACCAAGCTAGTGAATAGACCATCGGGCG 3639  
QY 4139 GTCAAGTAGTCAAGACCGCAACAAAATTTCACTGACAGGAACTTTTGAATCTTCA 4198  
DB 3640 GTCAAGTAGTCAAGACCGCAACAAAATTTCACTGACAGGAACTTTTGAATCTTCA 3699  
QY 4199 AAGTTGCTATTCAAGTAGTCAATTTGGGAGCATCAATATAGGGATTTATACGAGCAAC 4258  
DB 3700 AAGTTGCTATTCAAGTAGTCAATTTGGGAGCATCAATATAGGGATTTATACGAGCAAC 3759  
QY 4259 AGTGAAGTCAATCTACCAACTTTTGGCTCTCAGAAAAGCATAAACAGTTCTACTACC 4318  
DB 3760 AGTGAAGTCAATCTACCAACTTTTGGCTCTCAGAAAAGCATAAACAGTTCTACTACC 3819  
QY 4319 GCATTAGTGAACCTTTTCAATTCGCCAGTGGAGAAAAGCAAGCATCTAGC 4378  
DB 3820 GCATTAGTGAACCTTTTCAATTCGCCAGTGGAGAAAAGCAAGCATCTAGC 3879  
QY 4379 ATTAGCGGCAAGGATGCACTTTTATCAACAGGCTCTATAGATAACCCCTAGCGCTGG 4438  
DB 3880 ATTAGCGGCAAGGATGCACTTTTATCAACAGGCTCTCTATAGATAACCCCTAGCGCTGG 3939  
QY 4439 GATCATCTCTTTGGACAACCTTTTCTGCCAAATCTAGTCCAAATCACTTCATTGATACC 4498  
DB 3940 GATCATCTCTTTGGACAACCTTTTCTGCCAAATCTAGTCCAAATCACTTCATTGATACC 3999  
QY 4499 ATTATGTACAACTTGAGCAAGTGTGCGATCAGCTCCTCAAAATGGTCTCTGTAAACGGA 4558  
DB 4000 ATTAT- 4009  
QY 4559 TGACTCAACTTGCACTTAACCTTGAAGCTCAGTGCATTTAGTGAACCTTTGATCAGTTGTG 4618  
DB 4010 TGACTCAACTTGCACTTAACCTTGAAGCTCAGTGCATTTAGTGAACCTTTGATCAGTTGTG 4069  
QY 4619 CAGCTGGTCAGCAGATAGGAAAACACCGCTTTTCTTACCAAACTCAAGGAATTTATCAA 4678  
DB 4070 CAGCTGGTCAGCAGATAGGAAAACACCGCTTTTCTTACCAAACTCAAGGAATTTATCAA 4129  
QY 4679 CTCTCAACACCTTGGTATGCAAGTGTAGCAAGGGAATGTCTACTTTGAAGTCGGACAGTG 4738  
DB 4130 CTCTCAACACCTTGGTATGCAAGTGTAGCAAGGGAATGTCTACTTTGAAGTCGGACAGTG 4189  
QY 4739 AGTGTAGTCTTGAGAAATTTCTGAAGCGTATTTTATTTATCAGTGAGTCACTCAGGA 4798  
DB 4190 AGTGTAGTCTTGAGAAATTTCTGAAGCGTATTTTATTTATCAGTGAGTCACTCAGGA 4249

QY 4799 GATCCTCTAGCCGGAAGCATCGTGGCCGACCTGCAGGTGCGCATCACCGCGGCCACAGG 4858  
DB 4250 GATCCTCTAGCCGGAAGCATCGTGGC-----CGGCATCACCGCGGCCACAGG 4297  
QY 4859 TCGGTTTGTGGGCCCTTATATCCCGGACATCACGATGGGGAAGATCGGCGTGCCTCACTT 4918  
DB 4298 TCGGTTTGTGGGCCCTTATATCCCGGACATCACGATGGGGAAGATCGGCGTGCCTCACTT 4357  
QY 4919 CGGCTCATCAGGCGCTTTTTCGGCGTGGGTATGGTGGCAGGCCCCCGTGGCCGGGAGCT 4978  
DB 4358 CGGCTCATCAGGCGCTTTTTCGGCGTGGGTATGGTGGCAGGCCCCCGTGGCCGGGAGCT 4417  
QY 4979 GTTGGGCGCATCTCTTTGCAATGCCATTCCTTGGCGCGCGGTGCTCAAACGCGCTCAA 5038  
DB 4418 GTTGGGCGCATCTCTCTTGCATGCCATTCCTTGGCGCGCGGTGCTCAAACGCGCTCAA 4477  
QY 5039 CCTACTACTGGCTGCTTCTTAATCAGAGAGTCCGATAAGGAGAGCGTGCAGTATCTAT 5098  
DB 4478 CCTACTACTGGCTGCTTCTTAATCAGAGAGTCCGATAAGGAGAGCGTGCAGTATCTAT 4537  
QY 5099 GATTGGAAGTATGGGAATGGTATACCCGCACTTTCAGTGTCTTGAAGTCTCTATCAG 5158  
DB 4538 GATTGGAAGTATGGGAATGGTATACCCGCACTTTCAGTGTCTTGAAGTCTCTATCAG 4597  
QY 5159 ATTATGCCCAACTAAAGCAACCGGAGGAGAGATTTTCATGGTAAATTTCTCTGACTTTTG 5218  
DB 4598 ATTATGCCCAACTAAAGCAACCGGAGGAGAGATTTTCATGGTAAATTTCTCTGACTTTTG 4657  
QY 5219 GTCATCAGTAGACTCGAACTGTGAGACTATCTCCGTTATCAGACGAGAAATGTCCTTCTT 5278  
DB 4658 GTCATCAGTAGACTCGAACTGTGAGACTATCTCCGTTATCAGACGAGAAATGTCCTTCTT 4717  
QY 5279 GGAAGCAGTAAATGAAGTCCCAACAAATAAGAAATCTCTGTTATCAGGAACAACTTCTT 5338  
DB 4718 GGAAGCAGTAAATGAAGTCCCAACAAATAAGAAATCTCTGTTATCAGGAACAACTTCTT 4777  
QY 5339 GTTTCGAACCTTTTTCGGTGCCTTGAATATAAATATAGAGTGCATATGTCGGGTAGGAA 5398  
DB 4778 GTTTCGAACCTTTTTCGGTGCCTTGAATATAAATATAGAGTGCATATGTCGGGTAGGAA 4837  
QY 5399 TGAGAGGGGCAATGCTTACCTTCTGGAACCTTCAAGAGGATGATGAGGTTTGTAGTACT 5458  
DB 4838 TGAGAGGGGCAATGCTTACCTTCTGGAACCTTCAAGAGGATGATGAGGTTTGTAGTACT 4897  
QY 5459 GATGCCAATCTCAGTGACACGCTTATTTTCGTTCAAAACCTTCCGAATCCAGAGAAAT 5518  
DB 4898 GATGCCAATCTCAGTGACACGCTTATTTTCGTTCAAAACCTTCCGAATCCAGAGAAAT 4957  
QY 5519 CAAAGTTGTTTGTCTACTATTGATCCAAGCCAGTGGCGTCTTCAAACTGACAAATAGTGTG 5578  
DB 4958 CAAAGTTGTTTGTCTACTATTGATCCAAGCCAGTGGCGTCTTCAAACTGACAAATAGTGTG 5017  
QY 5579 CTGCTGTTTGTAGGTCATCTTTGATGAATAAATCTAGTCTTTGATCTAAATATCTTGA 5638  
DB 5018 CTGCTGTTTGTAGGTCATCTTTGATGAATAAATCTAGTCTTTGATCTAAATATCTTGA 5077  
QY 5639 CGAGCCAAAGGCGATAAATACCAAACTCTAAACCTCTTTTAAACCTTAAAGGACAAGTA 5698  
DB 5078 CGAGCCAAAGGCGATAAATACCAAACTCTAAACCTCTTTTAAACCTTAAAGGACAAGTA 5137  
QY 5699 TGTCTGCTGTATTAACCCCAAAATCAGCTCGTAGTCTGATCTCATCAACTTGTAGGGGC 5758  
DB 5138 TGTCTGCTGTATTAACCCCAAAATCAGCTCGTAGTCTGATCTCATCAACTTGTAGGGGC 5197  
QY 5759 ACTATCTTGTTTTGTAGGAAATTTTCGGAGATCGGATATCGAGAAAAGGTGACGCTGATTT 5818  
DB 5198 ACTATCTTGTTTTGTAGGAAATTTTCGGAGATCGGATATCGAGAAAAGGTGACGCTGATTT 5257  
QY 5819 TAAACGTGAAATTTATCTCAAGATCTCTGCCTCG 5852  
DB 5258 TAAACGTGAAATTTATCTCAAGATCGCGCCGCG 5291

## RESULT 15

US-10-179-046-19/c  
; Sequence 19, Application US/10179046  
; Publication No.: US20030013154A1  
; GENERAL INFORMATION:  
; APPLICANT: Crawford, Kenneth  
; Zaror, Isahel  
; Innis, Michael  
; TITLE OF INVENTION: Pichia Secretary Leader for Protein  
; Expression  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10179,046  
; FILING DATE: 25-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/029,267  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Chung, Ling-Fong  
; REGISTRATION NUMBER: 36,482  
; REFERENCE/DOCKET NUMBER: 1165.100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2704  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: nucleic acid  
; LENGTH: 2659 base pairs  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-10-179-046-19  
  
Query Match 32.3%; Score 2659; DB 15; Length 2659;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2145 GGATCTATCGAATCTAAATGTAAGTAAATCTCTAAATAATTAATAAGTCCAGTTTC 2204  
DB 2659 GGATCTATCGAATCTAAATGTAAGTAAATCTCTAAATAATTAATAAGTCCAGTTTC 2600  
  
QY 2205 TCCATACGAACTTAACAGCATGGGTGAGCATCTAGACCTTCAACAGAGCCAGATCC 2264  
DB 2599 TCCATACGAACTTAACAGCATGGGTGAGCATCTAGACCTTCAACAGAGCCAGATCC 2540  
  
QY 2265 ATCACTGCTTGGCCAAATATGTTTCAGTCCCTCAGGAGTTACGTCTTGTGAAGTGAATC 2324  
DB 2539 ATCACTGCTTGGCCAAATATGTTTCAGTCCCTCAGGAGTTACGTCTTGTGAAGTGAATC 2480  
  
QY 2325 TTCTGGAAGTTGCAAGTGTACTCCGTGTATTGACGGGCATATCCGTACGTTGGCAAA 2384  
DB 2479 TTCTGGAAGTTGCAAGTGTACTCCGTGTATTGACGGGCATATCCGTACGTTGGCAAA 2420  
  
QY 2385 GTGTGGTTGGTACCGGAGGAGTAATCTCCAACTCTCTCTGGAGAGTAGGCCAACCAAC 2444  
DB 2419 GTGTGGTTGGTACCGGAGGAGTAATCTCCAACTCTCTCTGGAGAGTAGGCCAACCAAC 2360  
  
QY 2445 ACAGATCCAGCGTGTACTTGTATCAACATAGAGAGCATCTCGATTGCGAGATC 2504

DB 2359 ACAGATCCAGCGTGTACTTGTATCAACATAAGAGAGCAATTCCTCGATTGCGAGATC 2300  
QY 2505 AAGTGTTCAGGAGCGTACTGATTGAGCAATTTCCAAAGCCCTGCTGTAGTTGCAACCGAT 2564  
DB 2299 AAGTGTTCAGGAGCGTACTGATTGAGCAATTTCCAAAGCCCTGCTGTAGTTGCAACCGAT 2240  
QY 2565 AGGGTTGTAGAGTGTGCAATACACATCTGCGTACAATTTCAACCCCTTGGCAACTGCAAGCT 2624  
DB 2239 AGGGTTGTAGAGTGTGCAATACACATCTGCGTACAATTTCAACCCCTTGGCAACTGCAAGCT 2180  
QY 2625 TGGTTGTGAACAGCATCTTCAATTTCTGCAAGCTCTTGTCTGTCTATATCGACGCAAC 2684  
DB 2179 TGGTTGTGAACAGCATCTTCAATTTCTGCAAGCTCTTGTCTGTCTATATCGACGCAAC 2120  
QY 2685 AGAATCACCTGGGAATCAATACCATGTTTCAGCTTGAACAGAGGTCTGAGGCAACGAAA 2744  
DB 2119 AGAATCACCTGGGAATCAATACCATGTTTCAGCTTGAACAGAGGTCTGAGGCAACGAAA 2060  
QY 2745 TCTGATCAGCGTATTTATCAGCAATACATAGAACTTTCAGAGGCCCCAGCAGGCAATGCA 2804  
DB 2059 TCTGATCAGCGTATTTATCAGCAATACATAGAACTTTCAGAGGCCCCAGCAGGCAATGCA 2000  
QY 2805 ATACTACACAGGCGTGTGTCATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGG 2864  
DB 1999 ATACTACACAGGCGTGTGTCATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGG 1940  
QY 2865 TTTCCTGGAACCAATATTTTGTCACTTAGGAACAGTTCTTGTTCGTAAGCCATAGCA 2924  
DB 1939 TTTCCTGGAACCAATATTTTGTCACTTAGGAACAGTTCTTGTTCGTAAGCCATAGCA 1880  
QY 2925 GCTACTGCTGGGCGCTCTCTAGCAGCATACATTAGCACCACTTGTGGGCAACG 2984  
DB 1879 GCTACTGCTGGGCGCTCTCTAGCAGCATACATTAGCACCACTTGTGGGCAACG 1820  
QY 2985 TAGATGACTTCTGGGGTAAGGGTACCATCTCTTAGGTGGAATGCAAAAACAATTTCT 3044  
DB 1819 TAGATGACTTCTGGGGTAAGGGTACCATCTCTTAGGTGGAATGCAAAAACAATTTCT 1760  
QY 3045 TTGCAACAGCAACTTTTGGCAGGAACACCCAGCATCAGGGAAGTGGGAAGCAGATTCG 3104  
DB 1759 TTGCAACAGCAACTTTTGGCAGGAACACCCAGCATCAGGGAAGTGGGAAGCAGATTCG 1700  
QY 3105 GTTCCACCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTGCAAAACAGAGCAGACT 3164  
DB 1699 GTTCCACCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTGCAAAACAGAGCAGACT 1640  
QY 3165 ACACAGGCGCAAGTCTCAACTTGCACAGCTCTCGTCTAGTTGAGCTTCATGGAATTTCCCTG 3224  
DB 1639 ACACAGGCGCAAGTCTCAACTTGCACAGCTCTCGTCTAGTTGAGCTTCATGGAATTTCCCTG 1580  
QY 3225 AGCTTATCTATAGAGAGATCAATGGCTCTCTTAACTGTTATCTGGCAATTCGATAAGTTCC 3284  
DB 1579 AGCTTATCTATAGAGAGATCAATGGCTCTCTTAACTGTTATCTGGCAATTCGATAAGTTCC 1520  
QY 3285 TCTGGGAAGAGAGCTCTTAACACAGAGTGTCTTAAAGCGACTCCATCAAACTTTGGCAGTT 3344  
DB 1519 TCTGGGAAGAGAGCTCTTAACACAGAGTGTCTTAAAGCGACTCCATCAAACTTTGGCAGTT 1460  
QY 3345 AGTTCTAAAAGGGCTTTGTACCAATTTTGAACAGCATTTGTGCAAAATTTGGTTGACTAAT 3404  
DB 1459 AGTTCTAAAAGGGCTTTGTACCAATTTTGAACAGCATTTGTGCAAAATTTGGTTGACTAAT 1400  
QY 3405 TCCATAATCTCTCCGTTTCTGGATAGGAGCGAGGAGCATCTTCAATTTCTCTGTAG 3464  
DB 1399 TCCATAATCTCTCCGTTTCTGGATAGGAGCGAGGAGCATCTTCAATTTCTGTAG 1340  
QY 3465 GAGGCTTTAGAAAAGCTCAATTTTGCACAAATTCACACTTTAGAGAGGACTTTCTTTA 3524  
DB 1339 GAGGCTTTAGAAAAGCTCAATTTTGCACAAATTCACACTTTAGAGAGGACTTTCTTTA 1280  
QY 3525 GGTTTGGAATCTCTTTAGAGTGTGTTCCCTGCTGATCTCCGCTTGGCATCTCTCTTTCCCT 3584  
DB 1279 GGTTTGGAATCTCTTTAGAGTGTGTTCCCTGCTGATCTCCGCTTGGCATCTCTCTTTCCCT 1220



```
QY 3585 CTAGTGACCTTTAGGACCTTCATATCCAGGTTTCTCTCCACCTCGTCCAAAGTCACACGG 3644
Db |||||
QY 1219 CTAGTGACCTTTAGGACCTTCATATCCAGGTTTCTCTCCACCTCGTCCAAAGTCACACGG 1160
Db |||||
QY 3645 TACTTGGCACATCTAATAATGCAAAATAAATAAGTCAGACATTTCCAGGCTATATCT 3704
Db |||||
QY 1159 TACTTGGCACATCTAATAATGCAAAATAAATAAGTCAGACATTTCCAGGCTATATCT 1100
QY 3705 TCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCCCTAAATTTAGCGTTCAACAAA 3764
Db |||||
QY 1099 TCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCCCTAAATTTAGCGTTCAACAAA 1040
QY 3765 ACTTCTCGTCAAAATAACCGTTTGGTATAGAAACCTTCTGGAGCATTTGCTCTTAGATCC 3824
Db |||||
QY 1039 ACTTCTCGTCAAAATAACCGTTTGGTATAGAAACCTTCTGGAGCATTTGCTCTTAGATCC 980
QY 3825 CACAGGTGGCTTCCATGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGTTCC 3884
Db |||||
QY 979 CACAAGTGGCTTCCATGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGTTCC 920
QY 3885 AAGTGACAGAAACCAACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCAGAA 3944
Db |||||
QY 919 AAGTGACAGAAACCAACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCAGAA 860
QY 3945 TCCAAATTCGATACCCAGCAACTTTTGAAGTGTCTCCAGATGTAGCACCTTTATACCACAAA 4004
Db |||||
QY 859 TCCAAATTCGATACCCAGCAACTTTTGAAGTGTCTCCAGATGTAGCACCTTTATACCACAAA 800
QY 4005 CCGTGACGACGAGATTTGGTAGACTCCAGTTTGTGCTTATAGCTTCGGATAGACTTT 4064
Db |||||
QY 799 CCGTGACGACGAGATTTGGTAGACTCCAGTTTGTGCTTATAGCTTCGGATAGACTTT 740
QY 4065 TTGGACGAGTACACCAAGGCCCAACAGTAATTTAGAAGTCAAGCCACCAAAAGTAGTGAAT 4124
Db |||||
QY 739 TTGGACGAGTACACCAAGGCCCAACAGTAATTTAGAAGTCAAGCCACCAAAAGTAGTGAAT 680
QY 4125 AGACCATCGGGGGTCACTAGTCAAAAGCGCAACAAATTTCACTGACAGGGAACCTTT 4184
Db |||||
QY 679 AGACCATCGGGGGTCACTAGTCAAAAGCGCAACAAATTTCACTGACAGGGAACCTTT 620
QY 4185 TTGACATCTTCAGAAAGTTTCGATTTCAGTAGTCAATTTGCCGAGCATCAATATGGGAT 4244
Db |||||
QY 619 TTGACATCTTCAGAAAGTTTCGATTTCAGTAGTCAATTTGCCGAGCATCAATATGGGAT 560
QY 4245 ATACCAAGAACCAAGTGAAGTCAATCTACAACTTTTGGGTCTCAGAAAAAGCATAA 4304
Db |||||
QY 559 ATACCAAGAACCAAGTGAAGTCAATCTACAACTTTTGGGTCTCAGAAAAAGCATAA 500
QY 4305 ACAGTTCTACTACCGCCATTAGTGAACCTTTTCAANTCGCCAGTGGAGAGAAAAGGC 4364
Db |||||
QY 499 ACAGTTCTACTACCGCCATTAGTGAACCTTTTCAANTCGCCAGTGGAGAGAAAAGGC 440
QY 4365 ACAGCGATACTAGCATTTAGCGGCAAGGATGCAACTTTTATCAACCGGGTCTCTATAGATA 4424
Db |||||
QY 439 ACAGCGATACTAGCATTTAGCGGCAAGGATGCAACTTTTATCAACCGGGTCTCTATAGATA 380
QY 4425 ACCTAGCGCCTGGGATCATCTTTGGACAACCTTTTCCCAATCTAGTCCAAAATC 4484
Db |||||
QY 379 ACCCTAGCGCCTGGGATCATCTTTTGGACAACCTTTTCCCAATCTAGTCCAAAATC 320
QY 4485 ACTTCATTGATACCATTTATTTGACAACTTTGAGCAAGTTGTGATCAGCTCTCAAAATGG 4544
Db |||||
QY 319 ACTTCATTGATACCATTTATTTGACAACTTTGAGCAAGTTGTGATCAGCTCTCAAAATGG 260
QY 4545 TCCTCTGTAAACGGATGACTCAACTTGGACATTTAACTTGAAGCTCAGTCAAGTGAAC 4604
Db |||||
QY 259 TCCTCTGTAAACGGATGACTCAACTTGGACATTTAACTTGAAGCTCAGTCAAGTGAAC 200
QY 4605 TTGATCAGGTTGTGAGCTGTGACGATAGGGAACACGGCTTTTCTACCAAACTC 4664
Db |||||
QY 199 TTGATCAGGTTGTGAGCTGTGACGATAGGGAACACGGCTTTTCTACCAAACTC 140
```

```
QY 4665 AAGGAATTATCAAACTCTGCAACACTTGCCTATTCAGGTAGCAGGGAATGTCACTTT 4724
Db |||||
QY 139 AAGGAATTATCAAACTCTGCAACACTTGCCTATTCAGGTAGCAGGGAATGTCACTTT 80
QY 4725 GAAGTCGACAGTCAGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTATTCAGTGA 4784
Db |||||
QY 79 GAAGTCGACAGTCAGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTATTCAGTGA 20
QY 4785 GTCAGTCATCAGGAGATCC 4803
Db |||||
QY 19 GTCAGTCATCAGGAGATCC 1
```

Search completed: September 24, 2004, 19:35:30  
Job time : 2991.74 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 18:06:49 ; Search time 352.227 Seconds  
(without alignments)  
12984.105 Million cell updates/sec

Title: US-10-030-390-3  
Perfect score: 841  
Sequence: 1 agatctaacatccaagacg.....aaataacagttattattcg 8241

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5142.8	62.4	8584	1	US-08-358-160-66
2	5130.8	62.3	8590	1	US-08-358-160-70
3	4565.2	55.4	8157	1	US-08-358-160-65
C 4	2659	32.3	2659	3	US-09-029-367-19
C 5	2310.8	28.0	34303	2	US-08-735-609-4
C 6	2310.8	28.0	34303	2	US-08-735-609-4
C 7	2310.8	28.0	34303	3	US-09-315-372-4
C 8	2310.8	28.0	34303	3	US-09-244-752-4
C 9	2310.8	28.0	34303	3	US-09-245-497-4
C 10	2310.8	28.0	34303	4	US-09-562-919-4
C 11	2302	27.9	6727	2	US-08-125-462-2
C 12	2302	27.9	6727	2	US-08-891-848-2
C 13	2302	27.9	6799	2	US-08-125-462-5
C 14	2302	27.9	6799	2	US-08-891-848-5
C 15	2302	27.9	8491	2	US-08-757-439-1
C 16	2301.4	27.9	11627	4	US-08-470-881-1
17	2299.6	27.9	5438	3	US-08-809-326A-25
18	2299.6	27.9	5438	4	US-09-689-914A-25
19	2299.6	27.9	5438	4	US-09-689-913A-25
20	2299.6	27.9	5438	4	US-09-689-916A-25
21	2299.6	27.9	5658	3	US-08-809-326A-10
22	2299.6	27.9	5658	4	US-09-689-914A-10
23	2299.6	27.9	5658	4	US-09-689-913A-10
24	2299.6	27.9	5825	4	US-09-689-916A-10
C 25	2299.6	27.9	5825	4	US-08-809-513A-7
C 26	2298.8	27.9	8068	4	US-09-301-593-27
C 27	2298.8	27.9	8068	4	US-09-301-593-35

Sequence 1, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 40, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 14, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 78, Appli  
Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-08-358-160-66  
; Sequence 66, Application US/08358160  
; Patent No. 5663143

; GENERAL INFORMATION:

; APPLICANT: LEY, Arthur C.

; APPLICANT: LADNER, Robert C.

; APPLICANT: GUTERMAN, Sonia K.

; APPLICANT: ROBERTS, Bruce L.

; APPLICANT: MARKLAND, William

; APPLICANT: KENT, Rachel B.

; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ

; NUMBER OF SEQUENCES: 234

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK

; STREET: 419 Seventh Street, N.W. Suite 300

; CITY: Washington

; STATE: District of Columbia

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/358,160

; FILING DATE: 15-DEC-1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/133,031

; FILING DATE: 13-OCT-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/009,319

; FILING DATE: 26-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/664,989

; FILING DATE: 01-MAR-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/487,063

; FILING DATE: 02-MAR-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/240,160

; FILING DATE: 02-SEP-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: Cooper, Iver P.

; REGISTRATION NUMBER: 28,005

; REFERENCE/DOCKET NUMBER: LEY=1

; TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8584 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: DNA plasmid  
US-08-358-160-66

Query Match 62.4%; Score 5142.8; DB 1; Length 8584;  
Best Local Similarity 94.6%; Pred. No. 0;  
Matches 5535; Conservative 0; Mismatches 167; Indels 152; Gaps 14;

QY 2 GATCTAACATCCAAAGACGAAAGTTGAATGAAACCTTTTGGCCATCCGACATCCACAGG 61  
DB 14 GATCTAACATCCAAAGACGAAAGTTGAATGAAACCTTTTGGCCATCCGACATCCACAGG 73  
QY 62 TCCATTCTCACATAGTGCMAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 121  
DB 74 TCCATTCTCACATAGTGCMAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 133  
QY 122 GCAACGCGAGACCTCCACTCTCTCTCTCAACACCCACTTTTGGCATCGAAACCA 181  
DB 134 GCAACGCGAGACCTCCACTCTCTCTCTCAACACCCACTTTTGGCATCGAAACCA 193  
QY 182 GCCAGATTATGGCTTGAATGGAGCTCGCTCTTCCAAATCTCTTATTAGGCTACTAA 241  
DB 194 GCCAGATTATGGCTTGAATGGAGCTCGCTCTTCCAAATCTCTTATTAGGCTACTAA 253  
QY 242 CACCATGACATTTATTAGCTGTCTATCTCGCCCCCTCGGAGGTTCAATGTTGTTAT 301  
DB 254 CACCATGACATTTATTAGCTGTCTATCTCGCCCCCTCGGAGG-TCAATGTTGTTAT 312  
QY 302 TTCGAATGCAACAGCTCCGATACACCGACATCACTCCAGATGAGGGCTTTCTGA 361  
DB 313 TTCGAATGCAACAGCTCCGATACACCGACATCACTCCAGATGAGGGCTTTCTGA 372  
QY 362 GTGTGGGTCAAATAGTTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTTAAACGCTG 421  
DB 373 GTGTGGGTCAAATAGTTTTCATGTT-CCCAAAATGGCCCAAACTGACAGTTTAAACGCTG 431  
QY 422 TCTTGGAACTTAATGACAAAGAGGTTGATCTATCCAGATGAACTAAAGTTTGGTTCTGT 481  
DB 432 TCTTGGAACTTAATGACAAAGAGGTTGATCTATCCAGATGAACTAAAGTTTGGTTCTGT 491  
QY 482 TGAATGCTAACGGCCAGTTCGTCGCAAAAGAACTTCCAAAGTCGCCATACCGTTTGTG 541  
DB 492 TGAATGCTAACGGCCAGTTCGTCGCAAAAGAACTTCCAAAGTCGCCATACCGTTTGTG 551  
QY 542 TTGTTTGGTATGATGACGAATGCTCAAAATTAATCTCATTAATGCTTAGCGCAGTCTC 601  
DB 552 TTGTTTGGTATGATGACGAATGCTCAAAATTAATCTCATTAATGCTTAGCGCAGTCTC 611  
QY 602 TCTATCGCTTCTGAACCCGGTGCACCTGTCGCAAAAGGCAAAATGGGGAAC-ACCGCT 660  
DB 612 TCTATCGCTTCTGAACCCGGTGCACCTGTCGCAAAAGGCAAAATGGGGAACCAACCGCT 671  
QY 661 TTTTGGATGATTAATGCAATGTT-CTCCACATGTTATGCTTCCAGATCTCGTGGGAATAC 719  
DB 672 TTTTGGATGATTAATGCAATGTTTCTCCACATGTTATGCTTCCAGATCTCGTGGGAATAC 731  
QY 720 TGCTGATAGCTTAAGTTCATGATCAAAATTAATGTTCTTAACCCCTACTTGACA-GCA 778  
DB 732 TGCTGATAGCTTAAGTTCATGATCAAAATTAATGTTCTTAACCCCTACTTGACGCA 791  
QY 779 ATATATAACAGAGGAAGCTGCCCTGTCTTTAAACCTTTTTTTTTTATCATCATTTATGAC 838  
DB 792 ATATATAACAGAGGAAGCTGCCCTGTCTTTAAACCTTTTTTTTTTATCATCATTTATGAC 851

QY 839 TTACTTTCAATAATTGCGACTGGTTCCAAATTGACAAGCTTTTGAATTTTACGACTTTTAAAC 898  
DB 852 TTACTTTCAATAATTGCGACTGGTTCCAAATTGACAAGCTTTTGAATTTTACGACTTTTAAAC 911  
QY 899 GACAACTTTGAGAAGATCAAAAAACAATAATTATTCGAAGGATCCAAACGATCAGATTTTC 958  
DB 912 GACAACTTTGAGAAGATCAAAAAACAATAATTATTCG-----AAACGATCAGATTTTC 963  
QY 959 CTTCAATTTTACTGCAAGTTTATTCGCAAGATCTCTCGCATTTAGCTGCTCCAGTCAACA 1018  
DB 964 CATCTATCTTCACTGCTGTTTTGTTGCTGCTCTCTGCTTGGTCTCCAGTTAAACA 1023  
QY 1019 CTACAACAGAGATGAACGCGCAAAATTCGGCTGAAGCTGTCATCGGTTACTTCAGATT 1078  
DB 1024 CCATCTACTGAAGACGAGCTGCTCAAAATTCCTGCTGAGGCTGTCATCGGTTACTCTGACT 1083  
QY 1079 TAGAAGGGGATTTTCGATGTTGCTGTTTGGCATTTTCCAAACAGCAAAATAACGGGTTAT 1138  
DB 1084 TGGAAAGGTGACTTCGACGCTGCTGTTTGGCAATCTCTAACTCTACTAACAACGGTTTGT 1143  
QY 1139 TGTTTTAAATACTACTATTGCGCAGCATTCGCTGAAGAGAGGGGTATCTCTCGAGA 1198  
DB 1144 TGTTCATCAACACTACCATTGCTTCTATCGCTGCTTAAGGAGGAAGGTTTCTCTGGACA 1203  
QY 1199 AAAGAGAGGCTGAAGCCCGAGGCCAGGCCAGGCCAGGAAAGAAACATGTATCA 1258  
DB 1204 AGAGAGCTGCTTGTAAAC----- 1220  
QY 1259 TGGCCCCCGGGAGAGAGATAAATTTGGCTTCCCGGTGTCCAGCCCGCAGCAGTGCACGG 1318  
DB 1221 TTGCCAAATCGTCAGAGGTCCATGCTTCTTTCCCAAGATGGGCTTTTCGAGC----- 1276  
QY 1319 AGAGAGGTTGCTGTTTGTGATGACAGTGTCCGGGATTCGGGTGCTTCCACCCCATGG 1378  
DB 1277 -----TGTTAAGGTTAAGTGGCTTTTCCCATACGGTGG 1312  
QY 1379 CCATCGAACAACCTCAAGAAAGAAATGTCCTTCTAACTAGTGGCGGTAGAAATCCCTAG 1438  
DB 1313 TTGTCAGAGTAAAC---GGTAAACAAGTTCTACTCTGAGAAGGAGGTAGAGAGTACTGTGG 1369  
QY 1439 GCGCGCCCGGATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTA 1498  
DB 1370 TGTTCATAGTAAGAAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCA-TTA 1428  
QY 1499 CGAGAAGACCGGCTCTGTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTGCCTGAG 1558  
DB 1429 CGAGAAGACCGGCTCTGTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTGCCTGAG 1488  
QY 1559 AGATCGAGCTTCATTTTGTATGACTTTTATTTGTAACCTATATAGTATAGGATTTTTT 1618  
DB 1489 AGATCGAGCTTCATTTTGTATGACTTTTATTTGTAACCTATATAGTATAGGATTTTTT 1548  
QY 1619 TTGTCATTTTGTCTTCTGTCGATCGAGCTTGTCTGATCAGCCTATCTCCAGCTGATG 1678  
DB 1549 TTGTCATTTTGTCTTCTGTCGATCGAGCTTGTCTGATCAGCCTATCTCCAGCTGATG 1608  
QY 1679 AATATCTTGTGTAGGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTGGTATTTTC 1738  
DB 1609 AATATCTTGTGTAGGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTGGTATTTTC 1668  
QY 1739 CCACCTCTCTCCAGAGTACAGAAATTAAGTGAGAAGTTTCGTTTGTGCAAGCTTATCGAT 1798  
DB 1669 CCACCTCTCTCCAGAGTACAGAAATTAAGTGAGAAGTTTCGTTTGTGCAAGCTTATCGAT 1728  
QY 1799 AAGCTTTAATGCGGTAGTTTATCAAGTTAATTCGTAACGAGTCCAGCACCCTGATG 1858  
DB 1729 AAGCTTTAATGCGGTAGTTTATCAAGTTAATTCGTAACGAGTCCAGCACCCTGATG 1789  
QY 1859 AATCTTAACAATGCGCTCATCGTCATCTCGCAGCGTCACTCGATGCTGTAGGCATA 1918  
DB 1789 AATCTTAACAATGCGCTCATCGTCATCTCGCAGCGTCACTCGATGCTGTAGGCATA 1848  
QY 1919 GGCTTGTTATGCGGTAAGTCCGGGCTCTTGGGGGATATCGTCCATTCGACAGCATC 1978

1849	Db	GGCTGTGGTTATGGCGTACTGCGGGGCTCTTGGGATATCGTCCATTCGACAGATC	1908
1979	Qy	GCAGTCACATATGGCGTGTCTAGCGCTATATGCGTGTGATGCAATTCCTATCGCACCC	2038
1909	Db	GCAGTCACATATGGCGTGTCTAGCGCTATATGCGTGTGATGCAATTCCTATCGCACCC	1968
2039	Qy	GTTCTCGAGCACTGTCCGACCGCTTTGGCGCGCGCCAGTCGTGTGCTTCGCTACTT	2098
1969	Db	GTTCTCGAGCACTGTCCGACCGCTTTGGCGCGCGCCAGTCGTGTGCTTCGCTACTT	2028
2099	Qy	GGAGCCACTATCGACTACGGCATCATGCGGACACACCCGCTCTGTGATCTATCGAATC	2158
2029	Db	GGAGCCACTATCGACTACGGCATCATGCGGACACACCCGCTCTGTGATCTATCGAATC	2088
2159	Qy	TAAATGTAAGTTAAATCTCTAAATAAATTAATAAGTCCAGTTTCCATACGAACCTT	2218
2089	Db	TAAATGTAAGTTAAATCTCTAAATAAATTAATAAGTCCAGTTTCCATACGAACCTT	2148
2219	Qy	AAAGCAATTCGGGTGACATCTAGACCTTCAAAGCAGCCAGATCCATCACTGCTTGGCC	2278
2149	Db	AAAGCAATTCGGGTGACATCTAGACCTTCAAAGCAGCCAGATCCATCACTGCTTGGCC	2208
2279	Qy	AAATATGTTTCAGTCCCTCAGAGTTACGTCTTGTGAAGTGAATCACTTCGGAAGTTGC	2338
2209	Db	AAATATGTTTCAGTCCCTCAGAGTTACGTCTTGTGAAGTGAATCACTTCGGAAGTTGC	2268
2339	Qy	AGTGTAACTCCGCTGTATTGAACGGGCAATATCCGTAGCTTGGCAAGTGTGTTGTACC	2398
2269	Db	AGTGTAACTCCGCTGTATTGAACGGGCAATATCCGTAGCTTGGCAAGTGTGTTGTACC	2328
2399	Qy	GGAGGAGTAATCTCCACAACTCTCTGAGAGTAGGCACCAACAAACACAGATCCAGCGTG	2458
2329	Db	GGAGGAGTAATCTCCACAACTCTCTGAGAGTAGGCACCAACAAACACAGATCCAGCGTG	2388
2459	Qy	TTGTACTTGATCAACATPAAGAAAGCAATTCCTCGATTTGCAAGTCAAGTGTTCAGAGC	2518
2389	Db	TTGTACTTGATCAACATPAAGAAAGCAATTCCTCGATTTGCAAGTCAAGTGTTCAGAGC	2448
2519	Qy	GTACTGATTTGGACATTTCCAAAGCCTCTCGTAGGTTGCAACCGATAGGTTGTAGAGTG	2578
2449	Db	GTACTGATTTGGACATTTCCAAAGCCTCTCGTAGGTTGCAACCGATAGGTTGTAGAGTG	2508
2579	Qy	TGCAATACACTTGGGTACAAATTTCAACCCCTTGGCAAATGCAACAGCTTGGTTGAAACGC	2638
2509	Db	TGCAATACACTTGGGTACAAATTTCAACCCCTTGGCAAATGCAACAGCTTGGTTGAAACGC	2568
2639	Qy	ATCTTCAATCTGCGAAGCTCTTGTGTATATCGACGCCAACAGATCACTCTGGGA	2698
2569	Db	ATCTTCAATCTGCGAAGCTCTTGTGTATATCGACGCCAACAGATCACTCTGGGA	2628
2699	Qy	ATCAATACCACTTGTGAAACCACTCTTGGCAGCAGTAGTAACAACTCTCGATCAGCGTA	2758
2629	Db	ATCAATACCACTTGTGAAACCACTCTTGGCAGCAGTAGTAACAACTCTCGATCAGCGTA	2687
2759	Qy	TTTATCAGCAATATCTAGAACTTCAGAAAGCCGACGAGGCAATGTCAATCTACACAGGC	2818
2688	Db	TTTATCAGCAATATCTAGAACTTCAGAAAGCCGACGAGGCAATGTCAATCTACACAGGC	2747
2819	Qy	TGATGTGTCAATTTTGAACCACTCTTGGCAGCAGTAGTAACAACTCTCTCGAACCAAA	2878
2748	Db	TGATGTGTCAATTTTGAACCACTCTTGGCAGCAGTAGTAACAACTCTCTCGAACCAAA	2807
2879	Qy	TATTTTGTCAACATTTAGAAACAGTTTCGTTCCGTAAAGCCATAGACGCTACTGCTGGCC	2938
2808	Db	TATTTTGTCAACATTTAGAAACAGTTTCGTTCCGTAAAGCCATAGACGCTACTGCTGGCC	2867
2939	Qy	GCCTCTGCTAGCAGATACACTTTAGCAACCACTTGTGGCAACGATGATGACTTCTGG	2998
2868	Db	GCCTCTGCTAGCAGATACACTTTAGCAACCACTTGTGGCAACGATGATGACTTCTGG	2927
2999	Qy	GGTAAGGGTACCCTCTTCTTAGGTGGAGATGCAAAAAACAATTTCTTTGCAACCAAGAAC	3058

D	b	2928	GGTAAGGGTACCAATCCCTTTAGGTGGAGATGCMAAAACAATTTCTTTTGCAACCCAGCAAC	2938
Q	y	3059	TTTGGCAGGAACACCCAGCATCAGGGAGTGGAGGCAGAAATTCGGTTTCCACACAGGAAT	3118
D	b	2988	TTTGGCAGGAACACCCAGCATCAGGGAAGTGGAGGCAGAAATTCGGTTTCCACACAGGAAT	3047
Q	y	3119	ATAGAGCCCAACTTTCTCAATAGGTCTTGCCAAACAGAGACGACATACACAGGGCAAGT	3178
D	b	3048	ATAGAGCCCAACTTTCTCAATAGGTCTTGCCAAACAGAGACGACATACACAGGGCAAGT	3107
Q	y	3179	CTCAACTTTGCAAAGCTCTCCGTTAGTTAGCTTTCATGGAAATTTCTCGTATTTCTATAGA	3238
D	b	3108	CTCAACTTTGCAAAGCTCTCCGTTAGTTAGCTTTCATGGAAATTTCTCGTATTTCTATAGA	3167
Q	y	3239	GAGATCAATGGCTCTTTAAGTTATCTGCAATTCGATAAGTTCTCTGGGAAAGGAGC	3298
D	b	3168	GAGATCAATGGCTCTTTAAGTTATCTGCAATTCGATAAGTTCTCTGGGAAAGGAGC	3227
Q	y	3299	TTCTAAACACAGGTGCTTTCAAAGCGACTCCATCAAACTTCGCGAGTTAGTTCTCTAAAGGGC	3358
D	b	3228	TTCTAAACACAGGTGCTTTCAAAGCGACTCCATCAAACTTCGCGAGTTAGTTCTCTAAAGGGC	3287
Q	y	3359	TTTGTCACCAATTTTGAGCAACATTTGTCGACAAATTTGGTTTGACTTAATTCCTAATTCCTGTC	3418
D	b	3288	TTTGTCACCAATTTTGAGCAACATTTGTCGACAAATTTGGTTTGACTTAATTCCTAATTCCTGTC	3347
Q	y	3419	CGTTTTCTGATAGGACGACGAGGGGCATCTTCAATTTCTTTGTGAGGAGCCCTTAGAAAC	3478
D	b	3348	CGTTTTCTGATAGGACGACGAGGGGCATCTTCAATTTCTTTGTGAGGAGCCCTTAGAAAC	3407
Q	y	3479	GTCAATTTTGCACAAATTCATACGACCTTCAGAGGGACTTCTTTAGTTTGGATTCCTTC	3538
D	b	3408	GTCAATTTTGCACAAATTCATACGACCTTCAGAGGGACTTCTTTAGTTTGGATTCCTTC	3467
Q	y	3539	TTTAGGTTTGTTCTTTGGTGATCTCTGGCTTTGGCATCTCTTCTCTCTAGTGACCTTTAG	3598
D	b	3468	TTTAGGTTTGTTCTTTGGTGATCTCTGGCTTTGGCATCTCTTCTCTCTAGTGACCTTTAG	3527
Q	y	3599	GGACTTCATATCCAGGTTTCTCTCCAGCTCGTCCAAAGTCACACCGTACTTGGCACATCT	3658
D	b	3528	GGACTTCATATCCAGGTTTCTCTCCAGCTCGTCCAAAGTCACACCGTACTTGGCACATCT	3587
Q	y	3659	AACTAATGCAAAATAAATAAGTCAGACATTTCCAGGCTATATCTTCTCTTGGATTTAGC	3718
D	b	3588	AACTAATGCAAAATAAATAAGTCAGACATTTCCAGGCTATATCTTCTCTTGGATTTAGC	3647
Q	y	3719	TTCTGCAAGTTTCATCAGCTTCTCCCTTAATTTTAGCGTTCAACAAAACTTCGTGCTCAAA	3778
D	b	3648	TTCTGCAAGTTTCATCAGCTTCTCCCTTAATTTTAGCGTTCAACAAAACTTCGTGCTCAAA	3707
Q	y	3779	TAAACGTTTGGTATAAGAACCTTCTGGAGCATTTGCTCTTAAGTCCCAAGGTTGGTTTC	3838
D	b	3708	TAAACGTTTGGTATAAGAACCTTCTGGAGCATTTGCTCTTAAGTCCCAAGGTTGGTTTC	3766
Q	y	3839	CATGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAAAGTCGGTTCCAGTGCACAGAAACC	3898
D	b	3767	CATGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAAAGTCGGTTCCAGTGCACAGAAACC	3826
Q	y	3899	AACACCTGTTTGTGTTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACC	3958
D	b	3827	AACACCTGTTTGTGTTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACC	3886
Q	y	3959	CAGCAACTTTTGTGTTGCTCAGATGTAGACCTTTTATACCAAAAACGTTGACACACGAGA	4018
D	b	3887	CAGCAACTTTTGTGTTGCTCAGATGTAGACCTTTTATACCAAAAACGTTGACACACGAGA	3946
Q	y	4019	TTGTTAGACTCCAGTTTGTGCTTTATAGCCTCTCCGGAATAGACTTTTGGACGAGTAGACAC	4078
D	b	3947	TTGTTAGACTCCAGTTTGTGCTTTTATAGCCTCTCCGGAATAGACTTTTGGACGAGTAGACAC	4006
Q	y	4079	CAGGCCCAACAGTAATTTAGAAAGTTCAGCCCAAAAAGTAGTGAATAGACCATCGGGGCG	4138
D	b	4007	CAGGCCCAACAGTAATTTAGAAAGTTCAGCCCAAAAAGTAGTGAATAGACCATCGGGGCG	4066

4139 GTCAAGTACCAAGACCCCAACAAATTTTCACTGACAGGGAATTTTGTGACATCTTTTCA 4198  
4067 GTCAAGTACCAAGACCCCAACAAATTTTCACTGACAGGGAATTTTGTGACATCTTTTCA 4126  
4199 AAGTTGTAATTCAGTAGTCAATTTGCCAGAGATCAATTAATGGGATTTATAGCAAGCAAC 4258  
4127 AAGTTGTAATTCAGTAGTCAATTTGCCAGAGATCAATTAATGGGATTTATAGCAAGCAAC 4186  
4259 AGTGGAAAGTCAATCTACCAACTTTTGGCGTCTCAGAAAAAGCATAAACAGTTCTACTACC 4318  
4187 AGTGGAAAGTCAATCTACCAACTTTTGGCGTCTCAGAAAAAGCATAAACAGTTCTACTACC 4246  
4319 GCCATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAAAGAAAGGACAGCATACTAGC 4378  
4247 GCCATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAAAGAAAGGACAGCATACTAGC 4306  
4379 ATTAGCGGCAAGATGCAACTTTTATCAACAGGCTCTATAGATACCTTAGCGCTGG 4438  
4307 ATTAGCGGCAAGATGCAACTTTTATCAACAGGCTCTATAGATACCTTAGCGCTGG 4366  
4439 GATCATCTTTTGGCAACTTTTCTGCAAAATCTAGTCCAAAATCACTTCAATGATACC 4498  
4367 GATCATCTTTTGGCAACTTTTCTGCAAAATCTAGTCCAAAATCACTTCAATGATACC 4426  
4499 ATTATTGTACAACTTTGAGCAAGTTGTGCATCAGCTCTCTCAATTTGGTCTCTCTAACGGA 4558  
4427 ATTATT-----ACGGA 4436  
4559 TGACTCAACTTTGCACATTAATCTGAAGCTCAGTGGATTGAGTGAACCTTGTACAGTTGTG 4618  
4437 TGACTCAACTTTGCACATTAATCTGAAGCTCAGTGGATTGAGTGAACCTTGTACAGTTGTG 4496  
4619 CAGCTGTGACAGATAGGAAACACAGGCTTTTCTTACCACCACTCAAGGAATTCACA 4678  
4497 CAGCTGTGACAGATAGGAAACACAGGCTTTTCTTACCACCACTCAAGGAATTCACA 4556  
4679 CTCTGCAACACTTTGCTATGAGTACAGTGGAAATGTCTATCTTGAAGTCCGACAGTG 4738  
4557 CTCTGCAACACTTTGCTATGAGTACAGTGGAAATGTCTATCTTGAAGTCCGACAGTG 4616  
4739 AGTGTAGTCTTGAAGAAATTCGAAAGCGGTATTTTATTTATCAGTGTAGTCAATCAGGA 4798  
4617 AGTGTAGTCTTGAAGAAATTCGAAAGCGGTATTTTATTTATCAGTGTAGTCAATCAGGA 4676  
4799 GATCTCTACCGCGAGCATCTGGCGGACCTGACGTCGAGTCCGATCACCGGCGCCACAGG 4858  
4677 GATCTCTACCGCGAGCATCTGGCG-----CGGCATCACCGGCGCCACAGG 4724  
4859 TCGGTTCTGGCGCTATATTCGCCGACATCACCGATGGGAAAGATCGGCTCGCCACTT 4918  
4725 TCGGTTCTGGCGCTATATTCGCCGACATCACCGATGGGAAAGATCGGCTCGCCACTT 4784  
4919 CGGCTCATAGCGCTTTTGGCGGTGGGTATGGTGGCGAGGCGCGGTGGCGCGGGGACT 4978  
4785 CGGCTCATAGCGCTTTTGGCGGTGGGTATGGTGGCGAGGCGCGGTGGCGCGGGGACT 4844  
4979 GTTGGGCGCCATCTCTTTCATGACATCCATTTCTTGGCGCGGCGGTGCTCAACGCGCTCAA 5038  
4845 GTTGGGCGCCATCTCTTTCATGACATCCATTTCTTGGCGCGGCGGTGCTCAACGCGCTCAA 4904  
5039 CCTACTAGTGGCTCTCTTTCATGAGGAGTCCNTAAGGAGAGCGTTCAGTATCTAT 5098  
4905 CCTACTAGTGGCTCTCTTTCATGAGGAGTCCNTAAGGAGAGCGTTCAGTATCTAT 4964  
5099 GATTGGAAGTATGGAAATGGTGTATACCGGCACTTCTTCAAGTGTCTTCAAGTCTCTCTATCAG 5158  
4965 GATTGGAAGTATGGAAATGGTGTATACCGGCACTTCTTCAAGTGTCTTCAAGTCTCTCTATCAG 5024  
5159 ATTATGCCCACTAAAGCAACCGGAGGAGGATTTTCAATGTAATTTCTGACTTTTG 5218  
5025 ATTATGCCCACTAAAGCAACCGGAGGAGGATTTTCAATGTAATTTCTGACTTTTG 5084

5219 GTCACTAGTACGTAAGTCCCAACCAATTAAGAAATCTTTGTTATCAGGAACAACTTTCTT 5278  
5085 GTCACTAGTACGTAAGTCCCAACCAATTAAGAAATCTTTGTTATCAGGAACAACTTTCTT 5144  
5279 GGAGACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTTGTTATCAGGAACAACTTTCTT 5338  
5145 GGAGACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTTGTTATCAGGAACAACTTTCTT 5204  
5339 GTTTCGAACTTTTTCGGTCCCTTGAACTATAAATATAGATGATATGTCGGGTAGGAA 5398  
5205 GTTTCGAACTTTTTCGGTCCCTTGAACTATAAATATAGATGATATGTCGGGTAGGAA 5264  
5399 TGGAGCGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGATACT 5458  
5265 TGGAGCGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGATACT 5324  
5459 GATGCCAACTTTTTCAGTGAACAGCTTCTATTTTCGTTTCAAACTTCGAAATCCAGAGAAAT 5518  
5325 GATGCCAACTTTTTCAGTGAACAGCTTCTATTTTCGTTTCAAACTTCGAAATCCAGAGAAAT 5384  
5519 CAAGTTCTTTTGTCTACTATTGATCCAGCCAGTCCGCTCTTGAACCTGACAAATAGTGTG 5578  
5385 CAAGTTCTTTTGTCTACTATTGATCCAGCCAGTCCGCTCTTGAACCTGACAAATAGTGTG 5444  
5579 CTCTGTTTTCAGGTCATCTTTTGTATGAATAAATCTAGTCTTTGATCTTAAATAATCTTGA 5638  
5445 CTCTGTTTTCAGGTCATCTTTTGTATGAATAAATCTAGTCTTTGATCTTAAATAATCTTGA 5504  
5639 CGAGCCAAAGGCGATTAATACCCAAATCTAAACCTTTTAAACGTTTAAAGGCAAGTA 5698  
5505 CGAGCCAAAGGCGATTAATACCCAAATCTAAACCTTTTAAACGTTTAAAGGCAAGTA 5564  
5699 TGCTGCTGCTGATTAACCCCAATCAGTCTGATCTGATCTCTCATCACTTCAGGCGGC 5758  
5585 TGCTGCTGCTGATTAACCCCAATCAGTCTGATCTGATCTCTCATCACTTCAGGCGGC 5624  
5759 ACTATCTTTTTCAGGAAATTTGCGGAGATGCGATATCGAGAAAGGATACGCTGATTT 5818  
5625 ACTATCTTTTTCAGGAAATTTGCGGAGATGCGATATCGAGAAAGGATACGCTGATTT 5684  
5819 TAAAGTGAATTTATCTCAAGATCTGCGCTCG 5852  
5685 TAAAGTGAATTTATCTCAAGATCGCGCGCGCG 5718

RESULT 2

US-08-358-160-70  
Sequence 70, Application US/08358160

Patent No. 563143

GENERAL INFORMATION:

APPLICANT: LEY, Arthur C.

APPLICANT: LADNER, Robert C.

APPLICANT: GUTERMAN, Sonia K.

APPLICANT: ROBERTS, Bruce L.

APPLICANT: MARKLAND, William

APPLICANT: KENT, Rachel B.

TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ

TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE

NUMBER OF SEQUENCES: 234

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W. Suite 300

CITY: Washington

STATE: District of Columbia

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/358,160

```
FILING DATE: 16-DEC-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/133,031
  FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/009,319
  FILING DATE: 26-JAN-1993
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/664,989
  FILING DATE: 01-MAR-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/487,063
  FILING DATE: 02-MAR-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/240,160
  FILING DATE: 02-SEP-1988
ATTORNEY/AGENT INFORMATION:
  NAME: Cooper, Iver P.
  REGISTRATION NUMBER: 28,005
  REFERENCE/DOCKET NUMBER: LEY=1
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 202-628-5197
    TELEFAX: 202-737-3528
    TELEX: 248633
  INFORMATION FOR SEQ ID NO: 70:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 8590 bases
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: circular
    MOLECULE TYPE: other nucleic acid
    DESCRIPTION: DNA plasmid
US-08-358-160-70

Query Match          62.3%; Score 5130.8; DB 1; Length 8590;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 5335; Conservative 0; Mismatches 167; Indels 154; Gaps 15;

QY 2 GATCTAACATCAAGAGCAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAGG 61
DB 14 GATCTAACATCAAGAGCAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAGG 73
QY 62 TCCATTCTCACATAAGTGCACACGCAAGAGGGGATACACTAGCAGCAGACCGTT 121
DB 74 TCCATTCTCACATAAGTGCACACGCAAGAGGGGATACACTAGCAGCAGACCGTT 133
QY 122 GCAAAACGAGGACCTCCACTCTCTCTCAACACCCCACTTTTGGCATCGAAAACCA 181
DB 134 GCAAAACGAGGACCTCCACTCTCTCTCAACACCCCACTTTTGGCATCGAAAACCA 193
QY 182 GCCCAGTTATTGGGCTTGAATGGAGCTCGCTCATTCGAAATTCCTTCTATTAGGCTACTAA 241
DB 194 GCCCAGTTATTGGGCTTGAATGGAGCTCGCTCATTCGAAATTCCTTCTATTAGGCTACTAA 253
QY 242 CACCATGACTTTATTAGCTGCTATCTCTGCGCCCTCGGAGGTTTCATGTTGTTTAT 301
DB 254 CACCATGACTTTATTAGCTGCTATCTCTGCGCCCTCGGAGG-TCATGTTGTTTAT 312
QY 302 TTCCGAATGCAACAGCTCCGATTAACCCGAAACATCACTCCAGATGAGGCTTTCTGA 361
DB 313 TTCCGAATGCAACAGCTCCGATTAACCCGAAACATCACTCCAGATGAGGCTTTCTGA 372
QY 362 GTGTGGGTCAAATAGTTTCATGTTTCCCAATGCCCCAAACTCACAGTTTAAACGCTG 421
DB 373 GTGTGGGTCAAATAGTTTCATGTTTCCCAATGCCCCAAACTCACAGTTTAAACGCTG 431
QY 422 TCTTGAACCTAATATGACAAAGCGTGATCTCATCCAAAGTAACTAAGTTGGTTGCT 481
DB 432 TCTTGAACCTAATATGACAAAGCGTGATCTCATCCAAAGTAACTAAGTTGGTTGCT 491
QY 482 TGAATGCTTAAGCGGCTTGGTCAAAAGAACTTCCAAAAGTCGCAATCCGCTTTCGTC 541
```





Db 3708 TAAACGTTGGTATGAAGACCTCTTGGAGCATGTCTTACGATCCCAAGGT-GCTTC 3766  
Qy 3839 CATGCTCTAAGACCCCTTTGATTGGCCAAAACAGGAAGTGGTCTCCAAAGTGACAGAAACC 3898  
Db 3767 CATGCTCTAAGACCCCTTTGATTGGCCAAAACAGGAAGTGGTCTCCAAAGTGACAGAAACC 3826  
Qy 3899 AACACCTGTTGTTCAACACAAATTCAGAGAGTCTCCATCAAAATCCAAATTCGATACC 3958  
Db 3827 AACACCTGTTGTTCAACACAAATTCAGAGAGTCTCCATCAAAATCCAAATTCGATACC 3886  
Qy 3959 CAGCAACTTTTCAGTGTGTCAGATGTAGCACCTTTATACCAAAACCGTGACACGAGA 4018  
Db 3887 CAGCAACTTTTCAGTGTGTCAGATGTAGCACCTTTATACCAAAACCGTGACACGAGA 3946  
Qy 4019 TTGTTAGACTCCAGTGTGTTGTCCTTATAGCTCCGGAATAGACTTTTGGACGAGTACAC 4078  
Db 3947 TTGTTAGACTCCAGTGTGTTGTCCTTATAGCTCCGGAATAGACTTTTGGACGAGTACAC 4006  
Qy 4079 CAGGCCCAACGAGTAATTAGAGAGTCTAGCCACCAAGTGTGTAATACCAATCCGGGCG 4138  
Db 4007 CAGGCCCAACGAGTAATTAGAGAGTCTAGCCACCAAGTGTGTAATACCAATCCGGGCG 4066  
Qy 4139 GTCACTAGTCAAGACGCGCAAAATTTCACTGACAGGGAACCTTTTGGACATCTTCAGA 4198  
Db 4067 GTCACTAGTCAAGACGCGCAAAATTTCACTGACAGGGAACCTTTTGGACATCTTCAGA 4126  
Qy 4199 AGTTTCGTATTCAGTAGTCAATTCGCGAGCATCAATATGGGATATACCAAGCAAC 4258  
Db 4127 AGTTTCGTATTCAGTAGTCAATTCGCGAGCATCAATATGGGATATACCAAGCAAC 4186  
Qy 4259 AGTGAAGTCACTACCACTTTGCGGTCTCAGAAAAGCATAAACAGTTCTACTACC 4318  
Db 4187 AGTGAAGTCACTACCACTTTGCGGTCTCAGAAAAGCATAAACAGTTCTACTACC 4246  
Qy 4319 GCATTAGTGAACCTTTCAATCCCGCAGTGGAGAGAAAAGGCAAGCGATCTAGC 4378  
Db 4247 GCATTAGTGAACCTTTCAATCCCGCAGTGGAGAGAAAAGGCAAGCGATCTAGC 4306  
Qy 4379 ATTACGGGCAAGGATGCAACTTTATCAACAGAGTCTCTATAGATAACCTTAGCGCTGG 4438  
Db 4307 ATTACGGGCAAGGATGCAACTTTATCAACAGAGTCTCTATAGATAACCTTAGCGCTGG 4366  
Qy 4439 GATCATCTTTGGACAACTTTCTTCGCAAAATCTAGGTCGAAATCACTTCATTGATACC 4498  
Db 4367 GATCATCTTTGGACAACTTTCTTCGCAAAATCTAGGTCGAAATCACTTCATTGATACC 4426  
Qy 4499 ATTATGTACAACTTGACCAAGTGTGTCATCAGCTCTCAATTTGGTCTCTGTAAACGA 4558  
Db 4427 ATTAT-----ACGGA 4436  
Qy 4559 TGACTCAACTTGCACTTAACCTTGAAGCTCAGTTCGATTGAGTGAACCTTGATCAGGTTGTG 4618  
Db 4437 TGACTCAACTTGCACTTAACCTTGAAGCTCAGTTCGATTGAGTGAACCTTGATCAGGTTGTG 4496  
Qy 4619 CAGCTGGTCAGCAGATAGGGAACACCGCTTTTCTTACCAAACTCAAGGAATATCAAA 4678  
Db 4497 CAGCTGGTCAGCAGATAGGGAACACCGCTTTTCTTACCAAACTCAAGGAATATCAAA 4556  
Qy 4679 CTCTGCAACACTTTGGTATGACAGTGTAGCAAGGGAATGTCTACTTTGAAGTCGACAGTG 4738  
Db 4557 CTCTGCAACACTTTGGTATGACAGTGTAGCAAGGGAATGTCTACTTTGAAGTCGACAGTG 4616  
Qy 4739 AGTGTAGTCTTGAGAAATTCGAACCGTATTTTATTTATCAGTGTAGTCACTCAAGGA 4798  
Db 4617 AGTGTAGTCTTGAGAAATTCGAACCGTATTTTATTTATTTATCAGTGTAGTCACTCAAGGA 4676  
Qy 4799 GATCCTCTACCGCGGACGATCGTGGCGGACCTGCGAGTCTCGGATCACCAGGCGCACAGG 4858  
Db 4677 GATCCTCTACCGCGGACGATCGTGGC-----CGGATCACCAGGCGCACAGG 4724  
Qy 4859 TCGGCTGTGCGGCTATATCGCGACATCACCGATCGGGAAGATCGGCTCGCCACTT 4918

Db 4725 TCGGTTGCTGGCGCCTATATCCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTT 4784  
Qy 4919 CCGGCTCATGAGCGCTTGTTTTCGGCCGTGGGTATGTTGGCAGGCCCCGTGGCCGGGGACT 4978  
Db 4785 CCGGCTCATGAGCGCTTGTTTTCGGCCGTGGGTATGTTGGCAGGCCCCGTGGCCGGGGACT 4844  
Qy 4979 GTTGGGCGCATCTCTCTGATGACCAATTCCTTGGCGGGCGGTGCTCAACGGCTCAA 5038  
Db 4845 GTTGGGCGCATCTCTCTGATGACCAATTCCTTGGCGGGCGGTGCTCAACGGCTCAA 4904  
Qy 5039 CTTACTACTGGGCTGCTTCTTAATGAGGAGTCCGATTAAGGAGAGCGTCAAGTATCTAT 5098  
Db 4905 CTTACTACTGGGCTGCTTCTTAATGAGGAGTCCGATTAAGGAGAGCGTCAAGTATCTAT 4964  
Qy 5099 GATTGGAAGTATGGGAATGTTGATACCCGATTCCTCAGTGCTTGGAGTCTCTATCAG 5158  
Db 4965 GATTGGAAGTATGGGAATGTTGATACCCGATTCCTCAGTGCTTGGAGTCTCTATCAG 5024  
Qy 5159 ATTATGCCAACTAAAGCAACCGGAGGAGAGATTCATGGTAAATTTCTCTGACTTTTG 5218  
Db 5025 ATTATGCCAACTAAAGCAACCGGAGGAGAGATTCATGGTAAATTTCTCTGACTTTTG 5084  
Qy 5219 GTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCT 5278  
Db 5085 GTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCT 5144  
Qy 5279 GAGACAGTAAATGAAGTCCCAATTAAGAAATCCTTGTATCAGGACAACTTCTT 5338  
Db 5145 GAGACAGTAAATGAAGTCCCAATTAAGAAATCCTTGTATCAGGACAACTTCTT 5204  
Qy 5339 GTTTT--CGAACTTTTTCGGTGCCTTGAACTATAAAATGTAGAGTGGATATGTCGGTAGG 5396  
Db 5205 GTTTTCGGAACCTTTTTCGGTGCCTTGAACTATAAAATGTAGAGTGGATATGTCGGTAGG 5264  
Qy 5397 AATGGAGCGGCAAAATGCTTACCTTGGACCTTCAAGAGTATGAGGGTTGTAGATA 5456  
Db 5265 AATGGAGCGGCAAAATGCTTACCTTGGACCTTCAAGAGTATGAGGGTTGTAGATA 5324  
Qy 5457 CTGATGCCAACTTCAGTGACCAAGTTCCTATTTCGTTTCAAAACCATTCGAAATCCAGAGAA 5516  
Db 5325 CTGATGCCAACTTCAGTGACCAAGTTCCTATTTCGTTTCAAAACCATTCGAAATCCAGAGAA 5384  
Qy 5517 ATCAAAGTGTGTTGTTCTATTTATGATCCAAAGCGAGTGGTCTTGAACCTGCAATAGTG 5576  
Db 5385 ATCAAAGTGTGTTGTTCTACTTATGATCCAAAGCGAGTGGTCTTGAACCTGCAATAGTG 5444  
Qy 5577 TGCTCGTGTGTTGAGGTCTCTTTGATGAATAAATCTAGTCTTTGATCTTAATAATCTT 5636  
Db 5445 TGCTCGTGTGTTGAGGTCTCTTTGATGAATAAATCTAGTCTTTGATCTTAATAATCTT 5504  
Qy 5637 GACGAGCCAGGCGATAAATACCCAAATCTAAACCTCTTTTAAACGTTTAAAGGACAG 5696  
Db 5505 GACGAGCCAGGCGATAAATACCCAAATCTAAACCTCTTTTAAACGTTTAAAGGACAG 5564  
Qy 5697 TATGCTGCTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCTCATCACTTGAAGG 5756  
Db 5565 TATGCTGCTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCTCATCACTTGAAGG 5624  
Qy 5757 GCACTATCTTGTGTTTGAAGAAATTTTCGGAGATGCGATATCGAGAAAAAGGTACGCTGAT 5816  
Db 5625 GCACTATCTTGTGTTTGAAGAAATTTTCGGAGATGCGATATCGAGAAAAAGGTACGCTGAT 5816  
Qy 5817 TTTAAACGTAATTTATCTCAAGATCTCTGCTCGCTG 5852  
Db 5685 TTTAAACGTAATTTATCTCAAGATCGCGCGCGG 5720

## RESULT 3

US-08-358-160-65  
; Sequence 65, Application US/08358160  
; Patent No. 5663143  
; GENERAL INFORMATION:  
; APPLICANT: LEV, Arthur C.

APPLICANT: LADNER, Robert C.  
APPLICANT: GUTERMAN, Sonia K.  
APPLICANT: ROBERTS, Bruce L.  
APPLICANT: MARKLAND, William  
APPLICANT: KENT, Rachel B.  
TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ  
TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE  
NUMBER OF SEQUENCES: 234  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W. Suite 300  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/358,160  
FILING DATE: 16-DEC-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/133,031  
FILING DATE: 13-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,319  
FILING DATE: 26-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/664,989  
FILING DATE: 01-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/487,063  
FILING DATE: 02-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/240,160  
FILING DATE: 02-SEP-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Cooper, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: LEV-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8157 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: CIRCULAR  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: DNA plasmid  
US-08-358-160-65

Query Match 55.4%; Score 4565.2; DB 1; Length 8157;  
Best Local Similarity 89.9%; Pred. No. 0;  
Matches 5262; Conservative 0; Mismatches 13; Indels 579; Gaps 11;

QY 2 GATCTAACATCCAAAGCAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 61  
DB 14 GATCTAACATCCAAAGCAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 73  
QY 62 TCCATCTCACACATAAGTGCACACGCAACAGGAGGGATACACTAGCAGCAGACCGTT 121  
DB 74 TCCATCTCACACATAAGTGCACACGCAACAGGAGGGATACACTAGCAGCAGACCGTT 133  
QY 122 GCAACGAGGAGGACCTCCACTCTCTCTCCCTCAACACCCACTTTTGGCCATCGAAAAACCA 181  
DB 134 GCAACGAGGAGGACCTCCACTCTCTCTCTCCCTCAACACCCACTTTTGGCCATCGAAAAACCA 193

182 GCCAGTTATTGGGCTTGATTGGAGCTCGCTCAATCCAAATTCCTTCTATTAGGCTACTAA 241  
194 GCCAGTTATTGGGCTTGATTGGAGCTCGCTCAATCCAAATTCCTTCTATTAGGCTACTAA 253  
242 CACATGACTTTATTAGGCTGCTCTATCTGCGCCCTGGCGAGGTTTCATGTTGTTTAT 301  
254 CACATGACTTTATTAGGCTGCTCTATCTGCGCCCTGGCGAGG-TCATGTTGTTTAT 312  
302 TTCGAATGCAACAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTCTGA 361  
313 TTCGAATGCAACAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTCTGA 372  
362 GTGTGGGTCAAATAGTTTCAATGTTCCCAATGGCCCAAACTGACAGTTTAAACCTG 421  
373 GTGTGGGTCAAATAGTTTCAATGTTCCCAATGGCCCAAACTGACAGTTTAAACCTG 431  
422 TCTTGGAACTTAATATGACAAAAGCGTGATCTCATCAAGATGAACCTAAGTTTGGTTCGT 481  
432 TCTTGGAACTTAATATGACAAAAGCGTGATCTCATCAAGATGAACCTAAGTTTGGTTCGT 491  
482 TGAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGCCATACCGTTTGTG 541  
492 TGAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGCCATACCGTTTGTG 551  
542 TTGTTTGGTATTGATTGACGAATGCTCAAAAATTAATCTCATTAACTAGTCAGCAGTCTC 601  
552 TTGTTTGGTATTGATTGACGAATGCTCAAAAATTAATCTCATTAACTAGTCAGCAGTCTC 611  
602 TCTATCGCTTCTGAACCCCGTGCACTGTGCCGAAAGCGAAATGGGGAAC-ACCGCT 660  
612 TCTATCGCTTCTGAACCCCGTGCGACCTGTGCCGAAACGCAAAATGGGGAACCAACCCGCT 671  
661 TTTTGGATGATTATGCTATGT-CTCCACATTTGATGCTTCCAGATTTCTGTGGGAATAC 719  
672 TTTTGGATGATTATGCTATGTCTCCACATTTGATGCTTCCAGATTTCTGTGGGAATAC 731  
720 TGCTGATAGCCTTAACGTTTCATGATCAAAAATTTAACTGTTCTAACCCCTACTTGACA-GCA 778  
732 TGCTGATAGCCTTAACGTTTCATGATCAAAAATTTAACTGTTCTAACCCCTACTTGACAGCA 791  
779 ATATATAACAGAGAGAGAGCTGCCCTGTCTTAAACCTTTTATCATCATTTATAGC 838  
792 ATATATAACAGAGAGAGAGCTGCCCTGTCTTAAACCTTTTATCATCATTTATAGC 851  
839 TTACTTTTCATATTGCGACTGGTTCCAATTCGACAGCTTTTGAATTTAACGACTTTTAAAC 898  
852 TTACTTTTCATATTGCGACTGGTTCCAATTCGACAGCTTTTGAATTTAACGACTTTTAAAC 911  
899 GACAACTTGAGAGATCAAAAACAACTTAATTTGAGAGGATCCAAACGATGAGATTTC 958  
912 GACAACTTGAGAGATCAAAAACAACTTAATTTGAGAGGATCCAAACGATGAGATTTC 953  
959 CTTCAATTTTACTGAGTTTATTTCGCGAGCATCTCCGCAATGCTGCTCCAGTCAACA 1018  
954 CTTCAATTTTACTGAGTTTATTTCGCGAGCATCTCCGCAATGCTGCTCCAGTCAACA 953  
1019 CTACACAGAGATGAAACGGCACAATAATTCGGGCTGAAGCTGTCTATCGGTTACTCAGATT 1078  
954 CTACACAGAGATGAAACGGCACAATAATTCGGGCTGAAGCTGTCTATCGGTTACTCAGATT 953  
1079 TAGAGGGGATTTCCGATGTTGCTGTTTGGCCATTTTCCACAGCACAATAACGGGTTAT 1138  
954 TAGAGGGGATTTCCGATGTTGCTGTTTGGCCATTTTCCACAGCACAATAACGGGTTAT 953  
1139 TGTTTATTAATACTACTATTGCCAGCATTTGCTGTAAAGAGAGAGGGGTATCTCTCGAGA 1198  
954 TGTTTATTAATACTACTATTGCCAGCATTTGCTGTAAAGAGAGAGGGGTATCTCTCGAGA 953  
1199 AAAGAGAGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGAGAGAAACATGTATCA 1258  
954 AAAGAGAGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGAGAGAAACATGTATCA 953  
1259 TGGCCCCCGGAGAGGATAAATTTGGCTTCCCGGTTTCCCGGTTTCCCGCCAGCAGTCAACGG 1318

Db 954 ----- 953  
Qy 1319 AGAGAGTTGCTGTTTGTGATGACAGTGTCCGGGATTCCTGGTGTCTTCCACCCCATGG 1378  
Db 954 ----- 953  
Qy 1379 CCATCGAGACACTCAAGNAGNAGTGTCCCTTCTTAAGTGTAGTGGCGTAGAATTCCTAG 1438  
Db 954 ----- 953  
Qy 1439 GCGCGCCGCGAATTAATTCGCTTAGACATGACTGTTCCCTCAGTTCAGTTGGGCACTTA 1498  
Db 954 -----AGGAATTCGCTTAGACATGACTGTTCCCTCAGTTCAGTTGGGCA-TTA 1001  
Qy 1499 CGAGAGACCGGCTCTGTAGATTCTAATCAAGAGGATGTGAGATGCCATTTGCCCTGAG 1558  
Db 1002 CGAGAGACCGGCTCTGTAGATTCTAATCAAGAGGATGTGAGATGCCATTTGCCCTGAG 1061  
Qy 1559 AGATCGAGGCTTCATTTTGTATACATTTTATTTTGTAACTATATAGTATAGGATTTTT 1618  
Db 1062 AGATCGAGGCTTCATTTTGTATACATTTTATTTTGTAACTATATAGTATAGGATTTTT 1121  
Qy 1619 TTGTCAATTTTGTCTTCCTGAGGCTTGTCTGATGAGCTTATCTGCAAGCTTATCGAT 1678  
Db 1122 TTGTCAATTTTGTCTTCCTGAGGCTTGTCTGATGAGCTTATCTGCAAGCTTATCGAT 1181  
Qy 1679 AATATCTTGTGTAGGGTTTGGGAATATCAATCGAGTTTGTATGCTTTTCTTGTATTTTC 1738  
Db 1182 AATATCTTGTGTAGGGTTTGGGAATATCAATCGAGTTTGTATGCTTTTCTTGTATTTTC 1241  
Qy 1739 CCATCTCTCTTTCAGAGTACAGAAGATTAAAGTGAAGTTTCTTTGTGCAAGCTTATCGAT 1798  
Db 1242 CCATCTCTCTTTCAGAGTACAGAAGATTAAAGTGAAGTTTCTTTGTGCAAGCTTATCGAT 1301  
Qy 1799 AAGCTTTAATCGGTAGTTTATCAGATTAAATGCTTAACGAGTACGACCGGTATG 1858  
Db 1302 AAGCTTTAATCGGTAGTTTATCAGATTAAATGCTTAACGAGTACGACCGGTATG 1361  
Qy 1859 AAATCTAAATCGCTCATCGTATCTCGGACCGTCAACCTGGATGCTGTAGGCATA 1918  
Db 1362 AAATCTAAATCGCTCATCGTATCTCGGACCGTCAACCTGGATGCTGTAGGCATA 1421  
Qy 1919 GCTTGTGTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGACAGCATC 1978  
Db 1422 GCTTGTGTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGACAGCATC 1481  
Qy 1979 GCCAGTCACTATGCGTCTGTAGCGCTATATGGTTGATGCAATTTCTATGCGACCC 2038  
Db 1482 GCCAGTCACTATGCGTCTGTAGCGCTATATGGTTGATGCAATTTCTATGCGACCC 1541  
Qy 2039 GTTCTCGAGCACTGTCCGACCGCTTTGGCGCGCCCGCAGTCTGCTTCTGCTACTT 2098  
Db 1542 GTTCTCGAGCACTGTCCGACCGCTTTGGCGCGCCCGCAGTCTGCTTCTGCTACTT 1601  
Qy 2099 GGAGCCACTATCGACTACGGATCATGCGGACCCACACCCGTCCTGTGGATCTATCGAATC 2158  
Db 1602 GGAGCCACTATCGACTACGGATCATGCGGACCCACACCCGTCCTGTGGATCTATCGAATC 1661  
Qy 2159 TAAATGTAAATCTTAAATTAATTAAGTCCCAAGTTTCTCCATACGAACCTT 2218  
Db 1662 TAAATGTAAATCTTAAATTAATTAAGTCCCAAGTTTCTCCATACGAACCTT 1721  
Qy 2219 AACAGCATTCGGGTAGCATCTAGACCTTCAACAGCAGCCAGATCCATCTGTTGGCC 2278  
Db 1722 AACAGCATTCGGGTAGCATCTAGACCTTCAACAGCAGCCAGATCCATCTGTTGGCC 1781  
Qy 2279 AATATGTTTCAGTCCCTCAGAGTTTACGTCGTGTGAGTGTAGTAACTTCTGGAAGTTGC 2338  
Db 1782 AATATGTTTCAGTCCCTCAGAGTTTACGTCGTGTGAGTGTAGTAACTTCTGGAAGTTGC 1841  
Qy 2339 AGTGTAACTCCGCTGTATTCAGCGGCATATCCGTATGCGAAGTGTGTTGATACC 2398  
Db 1841

Db 1842 AGTGTAACTCCGCTGTATTTGACGGGCATATCCGTAGCTTGGCAAGTGTGTTGTATACC 1901  
Qy 2399 GGAGGAGTAACTTCCAAACTCTCTCTGGAGTAGGCACCAACAAACACAGATCCAGCGTG 2458  
Db 1902 GGAGGAGTAACTTCCAAACTCTCTCTGGAGTAGGCACCAACAAACACAGATCCAGCGTG 1961  
Qy 2459 TTCTACTTGTATCAACATTAAGAGAGATTCGATTTTCAGGATCAAGTGTTCAGAGC 2518  
Db 1962 TTCTACTTGTATCAACATTAAGAGAGATTCGATTTTCAGGATCAAGTGTTCAGAGC 2021  
Qy 2519 GTACTGATTTGGACATTTCCAAAGCCCTGCTCGTAGTTTGAACCGATAGGTTGTAGAGTG 2578  
Db 2022 GTACTGATTTGGACATTTCCAAAGCCCTGCTCGTAGTTTGAACCGATAGGTTGTAGAGTG 2081  
Qy 2579 TGAATACACTTTCGCTGATTAATTTCAACCCCTTGGCAACTGCAAGCTTGTGTGAACAGC 2638  
Db 2082 TGAATACACTTTCGCTGATTAATTTCAACCCCTTGGCAACTGCAAGCTTGTGTGAACAGC 2141  
Qy 2639 ATCTTCAATTTCTGGCAAGCTTCTGCTGATATCGACAGCCAAACAGAACTCACCTGGGA 2698  
Db 2142 ATCTTCAATTTCTGGCAAGCTTCTGCTGATATCGACAGCCAAACAGAACTCACCTGGGA 2201  
Qy 2699 ATCAATACCATTTTCAGCTGTGAGACAGAAGTCTGAGGCAACGAAATCTGGATCAGCGTA 2758  
Db 2202 ATCAATACCATTTTCAGCTGTGAG- CAGAAGTCTGAGGCAACGAAATCTGGATCAGCGTA 2260  
Qy 2759 TTTTATCAGCAATAACTAGAACTTCAGAAGCCCGCAGGCGATGTCAATACTACACAGGCG 2818  
Db 2261 TTTTATCAGCAATAACTAGAACTTCAGAAGCCCGCAGGCGATGTCAATACTACACAGGCG 2320  
Qy 2819 TGAATGTCTCAATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGTTTCTCTGGACAAA 2878  
Db 2321 TGAATGTCTCAATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGTTTCTCTGGACAAA 2380  
Qy 2879 TATTTTGTACACTTAGGAACAGTTTCTTCTCGTAAAGCCATAGCAGCTACTGCTGGCG 2938  
Db 2381 TATTTTGTACACTTAGGAACAGTTTCTTCTCGTAAAGCCATAGCAGCTACTGCTGGCG 2440  
Qy 2939 GCCTCTGCTAGCACCATACATTAAGCAACCACTTGTGGCAACGATAGACTTCTCTGG 2998  
Db 2441 GCCTCTGCTAGCACCATACATTAAGCAACCACTTGTGGCAACGATAGACTTCTCTGG 2500  
Qy 2999 GGTAAAGGTACCATCTTCTTAGTGGAGATGCAAAAAAATTTCTTTCGCAACCAAGCAAC 3058  
Db 2501 GGTAAAGGTACCATCTTCTTAGTGGAGATGCAAAAAAATTTCTTTCGCAACCAAGCAAC 2560  
Qy 3059 TTTGGCAGAAACACCCAGCATCAGGAGTGAAGGAGCAATTTGGGTTCCACCAAGGAAT 3118  
Db 2561 TTTGGCAGAAACACCCAGCATCAGGAGTGAAGGAGCAATTTGGGTTCCACCAAGGAAT 2620  
Qy 3119 ATAGAGGCCAACTTCTCAATAGGTCTTGCAAAAACGAGCAGACTTACACAGGCGAAGT 3178  
Db 2621 ATAGAGGCCAACTTCTCAATAGGTCTTGCAAAAACGAGCAGACTTACACAGGCGAAGT 2680  
Qy 3179 CTCAACTTGCACCGTCTCGTTAGTTGAGCTTTCAGGATTTTCTGACGTTATCTATAGA 3238  
Db 2681 CTCAACTTGCACCGTCTCGTTAGTTGAGCTTTCAGGATTTTCTGACGTTATCTATAGA 2740  
Qy 3239 GAGATCAATGCGCTCTCTTAACTTATCTGGCAATTCGCAATTTCTCTGGGAAAGGAGC 3298  
Db 2741 GAGATCAATGCGCTCTCTTAACTTATCTGGCAATTCGCAATTTCTCTGGGAAAGGAGC 2800  
Qy 3299 TTTCTAACACAGGTGCTTCAAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCG 3358  
Db 2801 TTTCTAACACAGGTGCTTCAAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCG 2860  
Qy 3359 TTTGTCCACCATTTTGCAGCAATTTGTCGCAATTTGGTTTGTACTAATTCCTAATCTGTTTC 3418  
Db 2861 TTTGTCCACCATTTTGCAGCAATTTGTCGCAATTTGGTTTGTACTAATTCCTAATCTGTTTC 2920  
Qy 3419 CTTTTTCTCGATAGGACGAGGAGGCGATCTTCAATTTTCTGTGAGGAGGCGTTTGAAGAC 3478  
Db 2921 CTTTTTCTCGATAGGACGAGGAGGCGATCTTCAATTTTCTGTGAGGAGGCGTTTGAAGAC 2980

QY 3479 GTCAATTTTGGCAATTAATCAATGAGCCTTTCAGAGGGGACTCTTTAGTTGGATCTTC 3338  
DB 2981 GTCAATTTTGGCAATTAATCAATGAGCCTTTCAGAGGGGACTCTTTAGTTGGATCTTC 3040  
QY 3539 TTTAGGTTCTTCTTGGTGTATCTTGGTGTGCACTCTCTTTCTTCTAGTACCTTTAG 3598  
DB 3041 TTTAGGTTCTTCTTGGTGTATCTTGGTGTGCACTCTCTTTCTTCTAGTACCTTTAG 3100  
QY 3599 GGACTTCATATCCAGGTTTCTCTCACCTCTGTCMAAGTCACACCGTACTTTGGCACATCT 3658  
DB 3101 GGACTTCATATCCAGGTTTCTCTCACCTCTGTCMAAGTCACACCGTACTTTGGCACATCT 3160  
QY 3659 AACTAATGCAAAATAAATAAGTCAGCACATCTCCAGGCTATATCTTCTTGGATTTAGC 3718  
DB 3161 AACTAATGCAAAATAAATAAGTCAGCACATCTCCAGGCTATATCTTCTTGGATTTAGC 3220  
QY 3719 TTTCTGAAGTTCATCAGCTTCTCTCTAATTTTACGTTTCAACAAAATCTTCTGTCGCAAA 3778  
DB 3221 TTTCTGAAGTTCATCAGCTTCTCTCTAATTTTACGTTTCAACAAAATCTTCTGTCGCAAA 3280  
QY 3779 TAACCGTTTGGTATAGAACCTTCTGGAGCATCTCTTACGATCCCAAGGTTGGCTTC 3838  
DB 3281 TAACCGTTTGGTATAGAACCTTCTGGAGCATCTCTTACGATCCCAAGGTTGGCTTC 3339  
QY 3839 CATGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGTTCCTCAAGTGCAGAAAAC 3898  
DB 3340 CATGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGTTCCTCAAGTGCAGAAAAC 3399  
QY 3899 AACACCTGTTTGTCAACCAACAAATTTCAAGCAGTCTCCATCAACAAATTCGATACC 3958  
DB 3400 AACACCTGTTTGTCAACCAACAAATTTCAAGCAGTCTCCATCAACAAATTCGATACC 3459  
QY 3959 CAGCAACTTTTGAAGTGTCTCAGATGTAGACCTTTTATACCAAAACCGTGAAGACGAGA 4018  
DB 3460 CAGCAACTTTTGAAGTGTCTCAGATGTAGACCTTTTATACCAAAACCGTGAAGACGAGA 3519  
QY 4019 TTGGTAGACTCCAGTTTGTCTTATAGCTTCCGGAATAGCTTTTGGACAGTACAC 4078  
DB 3520 TTGGTAGACTCCAGTTTGTCTTATAGCTTCCGGAATAGCTTTTGGACAGTACAC 3579  
QY 4079 CAGGCCCAACAGTAAATAGAGAGTCAAGCCCAAAAGTGTAGTAAATAGACATCGGGGG 4138  
DB 3580 CAGGCCCAACAGTAAATAGAGAGTCAAGCCCAAAAGTGTAGTAAATAGACATCGGGGG 3639  
QY 4139 GTCAAGTCAAGAGCGCAACAAATTTTCACTGACAGGGAATCTTTTGGATCTTTCAGA 4198  
DB 3640 GTCAAGTCAAGAGCGCAACAAATTTTCACTGACAGGGAATCTTTTGGATCTTTCAGA 3699  
QY 4199 AAGTTCTGATTCAGTGTCAATTTGCGGAGCATCAATTAATGGGATATACCAAGAAC 4258  
DB 3700 AAGTTCTGATTCAGTGTCAATTTGCGGAGCATCAATTAATGGGATATATACCAAGAAC 3759  
QY 4259 AGTGAAGTCAATCTACCAATTTTGGGCTCTCAGAAAGAGCAATAACAGTCTTACTACC 4318  
DB 3760 AGTGAAGTCAATCTACCAATTTTGGGCTCTCAGAAAGAGCAATAACAGTCTTACTACC 3819  
QY 4319 GCCATTAGTGAACCTTTTCAATTCGCGAGCATCAATTAATGGGATATACCAAGAAC 4378  
DB 3820 GCCATTAGTGAACCTTTTCAATTCGCGAGCATCAATTAATGGGATATATACCAAGAAC 3879  
QY 4379 ATTACGGGCAAGGATCAACTTTTCAATTCGCGAGCATCAATTAATGGGATATACCAAGAAC 4438  
DB 3880 ATTACGGGCAAGGATCAACTTTTCAATTCGCGAGCATCAATTAATGGGATATACCAAGAAC 3939  
QY 4439 GATCATCTTTGGACAACTCTTTCTGCAATATCAGTGTCAAAATCAGTCTTACTATACC 4498  
DB 3940 GATCATCTTTGGACAACTCTTTCTGCAATATCAGTGTCAAAATCAGTCTTACTATACC 3999  
QY 4499 ATTAATGTACAACCTTTGAGCAAGTTGTGATCAGCTCTCAAAATTTGGTCTCTGTAAACGGA 4558  
DB 4000 ATTAATGTACAACCTTTGAGCAAGTTGTGATCAGCTCTCAAAATTTGGTCTCTGTAAACGGA 4009

QY 4559 TGAATCAACTTTGGCAATTAATCAATGAGCCTTTCAGAGGGGACTCTTTAGTTGGATCTTC 4618  
DB 4010 TGAATCAACTTTGGCAATTAATCAATGAGCCTTTCAGAGGGGACTCTTTAGTTGGATCTTC 4069  
QY 4619 CAGCTTGGTCAAGAGGAGGAAACAGGCTTTTCTTACCAAACTCAAGGAAATTTATCAAA 4678  
DB 4070 CAGCTTGGTCAAGAGGAGGAAACAGGCTTTTCTTACCAAACTCAAGGAAATTTATCAAA 4129  
QY 4679 CTCTGCAACACTTCTGCTATGAGGAGGAAATGTCATATCTTGAAGTCGACAGTG 4738  
DB 4130 CTCTGCAACACTTCTGCTATGAGGAGGAAATGTCATATCTTGAAGTCGACAGTG 4189  
QY 4739 AGTGTAGTCTTTGAGAAATTTCTGAAGCGGTATTTTATATATCAAGTGTGATCATCAGGA 4798  
DB 4190 AGTGTAGTCTTTGAGAAATTTCTGAAGCGGTATTTTATATATCAAGTGTGATCATCAGGA 4249  
QY 4799 GATCTCTACCGCGGAGCGATCTGTCGCGGACCTTGCAGAGTCTGGGATCACCGGCGCGACAGG 4858  
DB 4250 GATCTCTACCGCGGAGCGATCTGTCGCGGACCTTGCAGAGTCTGGGATCACCGGCGCGACAGG 4297  
QY 4859 TGGGTTGTCTGGCGCTTATATCGCGACATCACCGATGGGGAAGATCGGGCTCGCCACTT 4918  
DB 4298 TGGGTTGTCTGGCGCTTATATCGCGACATCACCGATGGGGAAGATCGGGCTCGCCACTT 4357  
QY 4919 CGGGCTCATGAGCGCTTGTCTTGGCGGTATGTTGGGAGGCGCGCTCGCGGGGACT 4978  
DB 4358 CGGGCTCATGAGCGCTTGTCTTGGCGGTATGTTGGGAGGCGCGCTCGCGGGGACT 4417  
QY 4979 GTTGGGCGCATCTCTCTTGCATGACCATCTTCTTGGCGGCGGCTCTCAAACGGCTCAA 5038  
DB 4418 GTTGGGCGCATCTCTCTTGCATGACCATCTTCTTGGCGGCGGCTCTCAAACGGCTCAA 4477  
QY 5039 CCTACTACTGGCTGTCTTCTTAATGAGGAGTGTGATTAAGGAGGAGCGTCTGATCTAT 5098  
DB 4478 CCTACTACTGGCTGTCTTCTTAATGAGGAGTGTGATTAAGGAGGAGCGTCTGATCTAT 4537  
QY 5099 GATTGGAAGTATGGGAATGTGTATACCGCATCTTTCAGTGTCTTGGCTCTCTATCAG 5158  
DB 4538 GATTGGAAGTATGGGAATGTGTATACCGCATCTTTCAGTGTCTTGGCTCTCTATCAG 4597  
QY 5159 ATTATGCCAACTAAAGCAACCGGAGGAGGAGATTTTCATGTTAAATTTCTGACTTTTG 5218  
DB 4598 ATTATGCCAACTAAAGCAACCGGAGGAGGAGATTTTCATGTTAAATTTCTGACTTTTG 4657  
QY 5219 GTCATCTAGTACTGCAACTGTGAGACTATCTCGGTTATGACAGCAAGAAATGTCTTCTT 5278  
DB 4658 GTCATCTAGTACTGCAACTGTGAGACTATCTCGGTTATGACAGCAAGAAATGTCTTCTT 4717  
QY 5279 GGAGACAGTAAATGAGTCTCCACCAATTAAGAAATCTCTTGTATATCAGGAAACAACTCTT 5338  
DB 4718 GGAGACAGTAAATGAGTCTCCACCAATTAAGAAATCTCTTGTATATCAGGAAACAACTCTT 4777  
QY 5339 GTTTCGAACTTTTTCGGTGTCTTGAATCTAAATCTAGAGTGGATATCTCGGTAGGAA 5398  
DB 4778 GTTTCGAACTTTTTCGGTGTCTTGAATCTAAATCTAGAGTGGATATCTCGGTAGGAA 4837  
QY 5399 TGGAGCGGCAAAATGCTTCTGCACTTCTGAGCTTCTGAGAGTATGTTGGGTTTGTATGAT 5458  
DB 4838 TGGAGCGGCAAAATGCTTCTGCACTTCTGAGCTTCTGAGAGTATGTTGGGTTTGTATGAT 4897  
QY 5459 GATGCCAACTTTTCAGTGTGACAGCTTGTCTATCTTCTTCAAAACCAATTCGAAAT 5518  
DB 4898 GATGCCAACTTTTCAGTGTGACAGCTTGTCTATCTTCTTCAAAACCAATTCGAAAT 4957  
QY 5519 CAAAGTTGTTGTCTACTATTTGATTCGAGCTGCGGCTTCTTGAACCTGAACTAGTGTG 5578  
DB 4958 CAAAGTTGTTGTCTACTATTTGATTCGAGCTGCGGCTTCTTGAACCTGAACTAGTGTG 5017  
QY 5579 CTCTGTCTTTGAGGTCATCTTGTATGAATAAATCTAGTCTTGTATCTTAAATTAATCTTGA 5638  
DB 5018 CTCTGTCTTTGAGGTCATCTTGTATGAATAAATCTAGTCTTGTATCTTAAATTAATCTTGA 5077  
QY 5639 CGAGCGAGGGAATAATACCAATCTTAAACTCTTAACTCTTAAACGCTTAAAGCAAGTA 5698



3405 TCCATATCTGTTCCGTTTCTGGATAGGACGAGGCGCATCTTCAATTTCTTGAG 3464  
 Db TCCATATCTGTTCCGTTTCTGGATAGGACGAGGCGCATCTTCAATTTCTTGAG 1340  
 3465 GAGGCTTTAGAAACGTCATTTTGCAAAATTAATACGACCTTCAGAGGAGCTTTT 3524  
 Db GAGGCTTTAGAAACGTCATTTTGCAAAATTAATACGACCTTCAGAGGAGCTTT 1280  
 3525 GGTTTGATTTCTTTAGTTGTTCTTGGTGTATCTTGGCTTGCATCTCTTCCCT 3584  
 Db GGTTTGATTTCTTTAGTTGTTCTTGGTGTATCTTGGCTTGCATCTCTTCCCT 1220  
 3585 CTAGTGACCTTTAGGACTTTCATATCCAGGTTTCTCTCCACCTCGTCCAAAGGTCAACCG 3644  
 Db CTAGTGACCTTTAGGACTTTCATATCCAGGTTTCTCTCCACCTCGTCCAAAGGTCAACCG 1160  
 3645 TACTTGCACTATTAATGCAAAATAAATAAGTTCAGCACTTCCAGGCTATATCT 3704  
 Db TACTTGCACTATTAATGCAAAATAAATAAGTTCAGCACTTCCAGGCTATATCT 1100  
 3705 TCCTTGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTAATTTAGGTTTCAACAA 3764  
 Db TCCTTGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTAATTTAGGTTTCAACAA 1040  
 1099 TCCTTGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTAATTTAGGTTTCAACAA 3824  
 Db TCCTTGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTAATTTAGGTTTCAACAA 980  
 3825 CACAAGTGGCTTCCATGGCTTCAAGCTTCTTGAATGGCCAAAGAGGCTTCC 3884  
 Db CACAAGTGGCTTCCATGGCTTCAAGCTTCTTGAATGGCCAAAGAGGCTTCC 920  
 3885 AAGTGACAGAAACCAACCTGTTTGTTCACACCAATTTCAAGCAGTCTCCATCACA 3944  
 Db AAGTGACAGAAACCAACCTGTTTGTTCACACCAATTTCAAGCAGTCTCCATCACA 860  
 3945 TCCAAATTCGATACCCAGCACTTTGATTTGCTCCAGATAGCACCTTTTATACCA 4004  
 Db TCCAAATTCGATACCCAGCACTTTGATTTGCTCCAGATAGCACCTTTTATACCA 800  
 4005 CCGTGACGACGAGATTTGATGACTTCCAGTTTGTCTCTTATAGCTTCCGGAATAGACTTT 4064  
 Db CCGTGACGACGAGATTTGATGACTTCCAGTTTGTCTCTTATAGCTTCCGGAATAGACTTT 740  
 4065 TTGACGAGTACACAGGCCCAACGAGTAATTAAGAGAGTCAGCCACCAAGTAGTAAT 4124  
 Db TTGACGAGTACACAGGCCCAACGAGTAATTAAGAGAGTCAGCCACCAAGTAGTAAT 680  
 4125 AGACATCGGGGGCTGAGTAGTCAAGAGCGCAACCAAAATTTCACTGACAGGAACTTT 4184  
 Db AGACATCGGGGGCTGAGTAGTCAAGAGCGCAACCAAAATTTCACTGACAGGAACTTT 620  
 4185 TTGACATCTTCAGAAAGTTTCTATTAGTAGTCAATTTCCGAGCATCAATTAATGGGAT 4244  
 Db TTGACATCTTCAGAAAGTTTCTATTAGTAGTCAATTTCCGAGCATCAATTAATGGGAT 560  
 4245 ATACAGAGCAACAGTGGAGTACATCTACGACTTGGCTTCCAGAAAGCATAA 4304  
 Db ATACAGAGCAACAGTGGAGTACATCTACGACTTGGCTTCCAGAAAGCATAA 500  
 4305 ACAGTTCTACTACGCCATTTAGTGAATTTTCAATCGCCAGTGGAGAGAAAGGC 4364  
 Db ACAGTTCTACTACGCCATTTAGTGAATTTTCAATCGCCAGTGGAGAGAAAGGC 440  
 4365 ACAGGATCTAGCATTTAGCGGCAAGGATGCACTTTATCAACCGGTCCTTAGATA 4424  
 Db ACAGGATCTAGCATTTAGCGGCAAGGATGCACTTTATCAACCGGTCCTTAGATA 380  
 4425 ACCCTAGCGCTGGGATCATCTTTTGGCAACTCTTTCTGCAAAATCTAGTCCAAATC 4484  
 Db ACCCTAGCGCTGGGATCATCTTTTGGCAACTCTTTCTGCAAAATCTAGTCCAAATC 320

4485 ACTTCATTGATACATTATTGTACAACTTGAGCAAGTTGTCGATCAGCTCCTCAATTTGG 4544  
 Db ACTTCATTGATACATTATTGTACAACTTGAGCAAGTTGTCGATCAGCTCCTCAATTTGG 260  
 4545 TCCTCTGTAAGGATGACTCAACTTGCACATTAACTTGAAGCTCAGTCGATTGAGTGAAC 4604  
 Db TCCTCTGTAAGGATGACTCAACTTGCACATTAACTTGAAGCTCAGTCGATTGAGTGAAC 200  
 4605 TTGATCAGGTTTGCAGCTGTCAGCAGCATAGGAAACACGCGTTTCTTACAAACTC 4664  
 Db TTGATCAGGTTTGCAGCTGTCAGCAGCATAGGAAACACGCGTTTCTTACAAACTC 140  
 4665 AAGGAATTCACAACTTGCACACTTGCATGTCAGTACGAGGAAATGTCATCTT 4724  
 Db AAGGAATTCACAACTTGCACACTTGCATGTCAGTACGAGGAAATGTCATCTT 80  
 4725 GAAGTCGACAGTGCAGTGTAGTCTTGAATAATCTGAAGCCGTTATTTATTCAGTGA 4784  
 Db GAAGTCGACAGTGCAGTGTAGTCTTGAATAATCTGAAGCCGTTATTTATTCAGTGA 20  
 4785 CTCAGTCATCAGGAGATCC 4803  
 Db CTCAGTCATCAGGAGATCC 1

RESULT 5

US-08-735-609-4/c  
 ; Sequence 4, Application US/08735609  
 ; Patent No. 5955360  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chamberlain, Jeffrey S.  
 ; APPLICANT: Amalfitano, Andrea  
 ; APPLICANT: Hauser, Michael A.  
 ; APPLICANT: Kumar-Singh, Rajendra  
 ; APPLICANT: Hartigan-O'Connor, Dennis J.  
 ; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS  
 ; NUMBER OF SEQUENCES: 15  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Medlen & Carroll, LLP  
 ; STREET: 220 Montgomery Street, Suite 2200  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States Of America  
 ; ZIP: 94104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/735,609  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ingolia, Diane E.  
 ; REGISTRATION NUMBER: 40,027  
 ; REFERENCE/DOCKET NUMBER: UN-02484  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 705-8410  
 ; TELEFAX: (415) 397-8338  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 34303 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: other nucleic acid  
 ; DESCRIPTION: /desc = "DNA"  
 ; US-08-735-609-4

Query Match 28.0%; Score 2310.8; DB 2; Length 34303;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 2356; Conservative 0; Mismatches 2; Indels 4; Gaps 1;



Qy	5845	CTGCCTCGCGGCTTTCCGGTGAATGACGGTGAATAAACCTCTGCACATNGCAGCTCCCGGAGAC	5904
Db	2846	CTGCCTCGCGGCTTTCCGGTGAATGACGGTGAATAAACCTCTGCACATNGCAGCTCCCGGAGAC	2787
Qy	5905	GGTCACAGCTTGTCTGTAAGCCGATGCCGGAGCAGACAAAGCCCGTCAGGCGCGGTGACG	5964
Db	2786	GGTCACAGCTTGTCTGTAAGCCGATGCCGGAGCAGACAAAGCCCGTCAGGCGCGGTGACG	2727
Qy	5965	GGGTGTTGGCGGTGTTCGGGGCGAGCCATATGACCCAGTCACTGATGCGATAGCGGAGTGTA	6024
Db	2726	GGGTGTTGGCGGTGTTCGGGGCGAGCCATATGACCCAGTCACTGATGCGATAGCGGAGTGTA	2667
Qy	6025	TACTGGCTTAACATATGCGGCATCAGAGCAGATTGACTGAGAGTGCACCAATATGCGGTGT	6084
Db	2666	TACTGGCTTAACATATGCGGCATCAGAGCAGATTGACTGAGAGTGCACCAATATGCGGTGT	2607
Qy	6085	GAATATCGGCACAGATCGTAAAGAGAAAAATACCGATACAGGCGCTCTTCGGCTTCCTCG	6144
Db	2606	GAATATCGGCACAGATCGTAAAGAGAAAAATACCGATACAGGCGCTCTTCGGCTTCCTCG	2547
Qy	6145	CTCACTGACTCGCTCGCTCGCTCGCTTCGGCTGCGGCGAGCGGTATCAGTCTCACTCAAAG	6204
Db	2546	CTCACTGACTCGCTCGCTCGCTTCGGCTGCGGCGAGCGGTATCAGTCTCACTCAAAG	2487
Qy	6205	CGCGTAAATACGGTTATCCACAGAAATACAGGGATTAACGCAGAAAGAACATGTGAGCAAAA	6264
Db	2486	CGCGTAAATACGGTTATCCACAGAAATACAGGGATTAACGCAGAAAGAACATGTGAGCAAAA	2427
Qy	6265	GGCCAGCAAAAGGCCAGGAAACCGTAAAAAAGCCGCTGTGCGGCTTTTCCATAGGCTC	6324
Db	2426	GGCCAGCAAAAGGCCAGGAAACCGTAAAAAAGCCGCTGTGCGGCTTTTCCATAGGCTC	2367
Qy	6325	CGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA	6384
Db	2366	CGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA	2307
Qy	6385	GGACTATAAGATACAGGCGTTTCCCCCTGGAGCTCCCTCGTGGCTCTCTCGTGTCCG	6444
Db	2306	GGACTATAAGATACAGGCGTTTCCCCCTGGAGCTCCCTCGTGGCTCTCTCGTGTCCG	2247
Qy	6445	ACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTTCT	6504
Db	2246	ACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTTCT	2187
Qy	6505	CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGTTGCCTCCAGCTGGGCTGT	6564
Db	2186	CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGTTGCCTCCAGCTGGGCTGT	2127
Qy	6565	GTGCACGAACCCCGCTTCAGCCCGACCGCTGGCGCTTATCCGGTAACATTCGTCATTGAG	6624
Db	2126	GTGCACGAACCCCGCTTCAGCCCGACCGCTGGCGCTTATCCGGTAACATTCGTCATTGAG	2067
Qy	6625	TCCAACCCGGTAAACACGACTTATCCGCACTGGCAGCAGCCACTGGTAAACAGATTAGC	6684
Db	2066	TCCAACCCGGTAAACACGACTTATCCGCACTGGCAGCAGCCACTGGTAAACAGATTAGC	2007
Qy	6685	AGACGAGGTATGTAGCGGTGTGTACAGAGTTCTTTGAAAGTGGTGGCCCTAACTACGGCTAC	6744
Db	2006	AGACGAGGTATGTAGCGGTGTGTACAGAGTTCTTTGAAAGTGGTGGCCCTAACTACGGCTAC	1947
Qy	6745	ACTAGAGGACAGTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	6804
Db	1946	ACTAGAGGACAGTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1887
Qy	6805	GTTGGTAGCTCTTGATCCGGGCAAAACCAACCGCTGGTACGGTGGTGTGTTGTTGTTGTTGTTG	6864
Db	1886	GTTGGTAGCTCTTGATCCGGGCAAAACCAACCGCTGGTACGGTGGTGTGTTGTTGTTGTTGTTG	1827
Qy	6865	AAGCAGCAGATTACCGCGCAAAAAAAGGATCTCAAGAGATCTCAAGAGATCTCTTTTCTACG	6924
Db	1826	AAGCAGCAGATTACCGCGCAAAAAAAGGATCTCAAGAGATCTCTTTTCTACG	1767

QY	6925	GGGTCTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGAGTTTGGTTCATGAGATTATCA	6984
DB	1766	GGGTCTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGAGTTTGGTTCATGAGATTATCA	1707
QY	6985	AAAAGGATCTTCACTAGATCCCTTTAAATATAAAATGAAGTTTAAATCAATCTAAAGT	7044
DB	1706	AAAAGGATCTTCACTAGATCCCTTTAAATATAAAATGAAGTTTAAATCAATCTAAAGT	1647
QY	7045	ATATATCAGTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGACCATATCTCA	7104
DB	1646	ATATATCAGTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGACCATATCTCA	1587
QY	7105	CGCATCTGTCTATTTCTGTTTCATCCATAGTTGCCTGACTCCCGTCTGTAGATAAATAAG	7164
DB	1586	CGCATCTGTCTATTTCTGTTTCATCCATAGTTGCCTGACTCCCGTCTGTAGATAAATAAG	1527
QY	7165	ATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCA	7224
DB	1526	ATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCA	1467
QY	7225	CCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAGGGCCGAGCGCAGAAAGTGGT	7284
DB	1466	CCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAGGGCCGAGCGCAGAAAGTGGT	1407
QY	7285	CCTGCAACTTTATCCGGCTCCATCCAGTCTAATTAATTGTTCCCGGAAGCTAGAGTAAGT	7344
DB	1406	CCTGCAACTTTATCCGGCTCCATCCAGTCTAATTAATTGTTCCCGGAAGCTAGAGTAAGT	1347
QY	7345	AGTTTCGCAGTTAATAGTTTTCGCGAACAGTTGTTGCCATGCTCGACGCATCGTGGTGTCA	7404
DB	1346	AGTTTCGCAGTTAATAGTTTTCGCGAACAGTTGTTGCCATGCTCGACGCATCGTGGTGTCA	1287
QY	7405	CGCTCGTGTGTTGGTATGGCTTCATTACAGCTCCGGTTCCCAAGCATCAAGCCAGATTACA	7464
DB	1286	CGCTCGTGTGTTGGTATGGCTTCATTACAGCTCCGGTTCCCAAGCATCAAGCCAGATTACA	1227
QY	7465	TGATCCCCCATGTTGTGCAAAAAAGCGGTTAGTCTCTTCGGTCCCTCCGATCGTTGTGAGA	7524
DB	1226	TGATCCCCCATGTTGTGCAAAAAAGCGGTTAGTCTCTTCGGTCCCTCCGATCGTTGTGAGA	1167
QY	7525	AGTAAGTTGCCCGAGTTATCAGTCATGTTTATGCGCAGCACTGCATAAATCTCTTACT	7584
DB	1166	AGTAAGTTGCCCGAGTTATCAGTCATGTTTATGCGCAGCACTGCATAAATCTCTTACT	1107
QY	7585	GTCA TGCCATCCGTAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGTCATTCTGA	7644
DB	1106	GTCA TGCCATCCGTAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGTCATTCTGA	1047
QY	7645	GAATAGTGTTATCGGCGACCGAGTTGCTCTTGCCCGCGTCAACACGGGATATACCGCG	7704
DB	1046	GAATAGTGTTATCGGCGACCGAGTTGCTCTTGCCCGCGTCAACACGGGATATACCGCG	987
QY	7705	CCATAGCAGAACTTTAAAAGTGTCTCATTTGGAAAAAGTTCCTTCGGGGCGAAAACTC	7764
DB	986	CCATAGCAGAACTTTAAAAGTGTCTCATTTGGAAAAAGTTCCTTCGGGGCGAAAACTC	927
QY	7765	TCAAGNATTTACGCTGTGTGAGATCAAGTTTGATTAACCACTCGTGCACCCAACTGA	7824
DB	926	TCAAGNATTTACGCTGTGTGAGATCAAGTTTGATTAACCACTCGTGCACCCAACTGA	867
QY	7825	TCCTTCAGCATCTTTTACTTTTCCACAGGTTTCTGGGTGAGCAAAAACAGAGGCAAAAT	7884
DB	866	TCCTTCAGCATCTTTTACTTTTCCACAGGTTTCTGGGTGAGCAAAAACAGAGGCAAAAT	807
QY	7885	CCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATCTCTTCCTTTT	7944
DB	806	CCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATCTCTTCCTTTT	747
QY	7945	CAATATATTGAAGCATTTATCAGGGTTATTGCTCATGACGGGATACATATTGAAATG	8004
DB	746	CAATATATTGAAGCATTTATCAGGGTTATTGCTCATGACGGGATACATATTGAAATG	687
QY	8005	ATTAGAAAAATAAAACAAATFAGGGTTCCGGCGACATTTTCCCGGAAAAAGTGCCACCTGAC	8064



QY 7105 GCGATCTGCTATTTTCGTTTCATCATAGTTCGCTGACATCCCGCTGCTAGATACTACG 7164  
Db 1586 GCGATCTGCTATTTTCGTTTCATCATAGTTCGCTGACATCCCGCTGCTAGATACTACG 1527  
QY 7165 ATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATACCGCGAGACCCAGCTCA 7224  
Db 1526 ATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATACCGCGAGACCCAGCTCA 1467  
QY 7225 CCGGCTCCAGATTTATCAGCAATAACAGCAGCCGGAAGGGCGAGCGAGAGTGGT 7284  
Db 1466 CCGGCTCCAGATTTATCAGCAATAACAGCAGCCGGAAGGGCGAGCGAGAGTGGT 1407  
QY 7285 CCGTCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAGAGTAAGT 7344  
Db 1406 CCGTCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAGAGTAAGT 1347  
QY 7345 AGTTCGCCAGTAATAGTTTCGGMAACGTTGTGTCATTCGTGAGGATCGTGGTGTCA 7404  
Db 1346 AGTTCGCCAGTAATAGTTTCGGMAACGTTGTGTCATTCGTGAGGATCGTGGTGTCA 1287  
QY 7405 CGCTCGTGGTTGGTATGCTTCATTGACGTCGCTTCCCAACGATCAAGCGAGTTACA 7464  
Db 1286 CGCTCGTGGTTGGTATGCTTCATTGACGTCGCTTCCCAACGATCAAGCGAGTTACA 1227  
QY 7465 TGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTCTCGTCTCTCCGATCGTTGTGCA 7524  
Db 1226 TGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTCTCGTCTCTCCGATCGTTGTGCA 1167  
QY 7525 AGTAGTTGGCGGAGTCTTACTCATGTTATGCGAGCACTGCATATTCCTTACT 7584  
Db 1166 AGTAGTTGGCGGAGTCTTACTCATGTTATGCGAGCACTGCATATTCCTTACT 1107  
QY 7585 GTCATGCCATCCGTAAGTGTCTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTGA 7644  
Db 1106 GTCATGCCATCCGTAAGTGTCTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTGA 1047  
QY 7645 GAATAGTGTATGCGCGACCGAGTGTCTTTCGCGGGGTCAACCGGATATACCGCG 7704  
Db 1046 GAATAGTGTATGCGCGACCGAGTGTCTTTCGCGGGGTCAACCGGATATACCGCG 987  
QY 7705 CCATACAGCAACTTTTAAAGTGTCTCATCTTGGMAAACGTTCTCGGGCGGAAACTC 7764  
Db 986 CCATACAGCAACTTTTAAAGTGTCTCATCTTGGMAAACGTTCTCGGGCGGAAACTC 927  
QY 7765 TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCACTCGTGACCCCACTGA 7824  
Db 926 TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCACTCGTGACCCCACTGA 867  
QY 7825 TCTTCAGCATCTTTTACTTTTCCAGCGTTTCTCGGTGAGCAAAACGAGGACCAAT 7884  
Db 866 TCTTCAGCATCTTTTACTTTTCCAGCGTTTCTCGGTGAGCAAAACGAGGACCAAT 807  
QY 7885 GCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATCTACTCTTCTCTTTT 7944  
Db 806 GCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATCTACTCTTCTCTTTT 747  
QY 7945 CAATATTTAGAGCTTTATCAGGTTATTTGTTCTCATGAGCGGATACATATTTGAATGT 8004  
Db 746 CAATATTTAGAGCTTTATCAGGTTATTTGTTCTCATGAGCGGATACATATTTGAATGT 687  
QY 8005 ATTTAGAAAATAAACAATAAGGGGTTCGCGGACATTTCCCGGAAAGTGCCACCTGAC 8064  
Db 686 ATTTAGAAAATAAACAATAAGGGGTTCGCGGACATTTCCCGGAAAGTGCCACCTGAC 627  
QY 8065 GTCTAAGAAACCATTTATCATGACATTAACCTATAAAAATAGCGGTATCACAGGCC 8124  
Db 626 GTCTAAGAAACCATTTATCATGACATTAACCTATAAAAATAGCGGTATCACAGGCC 567  
QY 8125 TTTGCTCTTCAAGATTAATTTCTCATGTTTTCAGAGCTTATCATCGATAAGCT 8176  
Db 566 TTTGCTCTTCAAG-----AATTTCTATGTTTTCAGAGCTTATCATCGATAAGCT 519

RESULT 7  
US-09-315-372-4/c  
; Sequence 4, Application US/09315372  
; Patent No. 6057158  
; GENERAL INFORMATION:  
; APPLICANT: Chamberlain, Jeffrey S.  
; APPLICANT: Amalfitano, Andrea  
; APPLICANT: Hauser, Michael A.  
; APPLICANT: Kumar-Singh, Rajendra  
; APPLICANT: Hartigan-O'Connor, Dennis J.  
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/315,372  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/735,609  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: UM-02484  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 34303 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-09-315-372-4

Query Match 28.0%; Score 2310.8; DB 3; Length 34303;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2326; Conservative 0; Mismatches 2; Indels 4; Gaps 1;  
QY 5845 CTGCCTCGCGCGTTTCGGTGATGACGGTGAAAAACCTCTGACACATGCAGCTCCCGGAGAC 5904  
Db 2846 CTGCCTCGCGCGTTTCGGTGATGACGGTGAAAAACCTCTGACACATGCAGCTCCCGGAGAC 2787  
QY 5905 GGTACAGCTTGTCTGTAAGCGGATGCCCGGAGAGACAAGCCCGTCAGGGCGGTCAGC 5964  
Db 2786 GGTACAGCTTGTCTGTAAGCGGATGCCCGGAGAGACAAGCCCGTCAGGGCGGTCAGC 2727  
QY 5965 GGGTGTTCGGGGTGTTCGGGGCGAGCCATGACCCAGTCACTAGCGATAGCGGAGTGA 6024  
Db 2726 GGGTGTTCGGGGTGTTCGGGGCGAGCCATGACCCAGTCACTAGCGATAGCGGAGTGA 2667  
QY 6025 TACTGCTTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATCGGTTG 6084  
Db 2666 TACTGCTTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATCGGTTG 2607  
QY 6085 GAAATACCCAGAGATCGGTAGAGGAAATACCCGATCAGCGCTCTTCCGCTTCCCTCG 6144  
Db 2606 GAAATACCCAGAGATCGGTAGAGGAAATACCCGATCAGCGCTCTTCCGCTTCCCTCG 2547  
QY 6145 CTCACTGACTCGTCTCGGCTCGGCTGTTGGCTGCGGCGAGCGGTATCAGCTCACTCAAG 6204



TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medlen & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States Of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/244,752  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/735,609  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ingolia, Diane E.  
REGISTRATION NUMBER: 40,027  
REFERENCE/DOCKET NUMBER: UM-02484  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34303 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "DNA"  
US-09-244-752-4

Query Match 28.0%; Score 2310.8; DB 3; Length 34303;

Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2326; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY	5845	CTGCTCCGCGGTTTCGGTGTATGACGGTGAAACCTCTGACACATGACGAGTCCCGGAGAC	5904
DB	2846	CTGCTCCGCGGTTTCGGTGTATGACGGTGAAACCTCTGACACATGACGAGTCCCGGAGAC	2787
QY	5905	GGTCACAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAGCCCGTCAGGCGCGTCAGC	5964
DB	2786	GGTCACAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAGCCCGTCAGGCGCGTCAGC	2727
QY	5965	GGGTGTTGGCGGGTGTGCGGGCGCAGCATGACCCAGTCAGTAGCGATAGCGGAGTGA	6024
DB	2726	GGGTGTTGGCGGGTGTGCGGGCGCAGCATGACCCAGTCAGTAGCGATAGCGGAGTGA	2667
QY	6025	TACTGGCTTAATATGCGGCATCAGACAGATTGTACTGAGAGTGACCATATGCGGTGT	6084
DB	2666	TACTGGCTTAATATGCGGCATCAGACAGATTGTACTGAGAGTGACCATATGCGGTGT	2607
QY	6085	GAATACCGCACAGATGCGTAAGAGAAATACCGCATCAGCGCTCTTCGGCTTCCTCG	6144
DB	2606	GAATACCGCACAGATGCGTAAGAGAAATACCGCATCAGCGCTCTTCGGCTTCCTCG	2547
QY	6145	CTCACTGACTCGCTCGGCTCGGTCGGTGGCGGAGCGGATCAGCTCACTCAAAAG	6204
DB	2546	CTCACTGACTCGCTCGGCTCGGTCGGTGGCGGAGCGGATCAGCTCACTCAAAAG	2487
QY	6205	CGGTATATACGGTTATCCACAGATCAGGGGATTAACGAGAAAGAACATGTAGCAAAA	6264
DB	2486	CGGTATATACGGTTATCCACAGATCAGGGGATTAACGAGAAAGAACATGTAGCAAAA	2427
QY	6265	GGCCAGCAAAAGGCCAGAACCGTAAAGAGCCGCTGCTGGCTTTTCCATAGGCTC	6324
DB	2426	GGCCAGCAAAAGGCCAGAACCGTAAAGAGCCGCTGCTGGCTTTTCCATAGGCTC	2367

QY	6325	CGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGCGGAACCCGACA	6384
DB	2366	CGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGCGGAACCCGACA	2307
QY	6385	GGACTATAAAGATACCAGCGGTTTCCCTCTGGAAGCTCCCTCTGCGCTCTCTCTGTCCG	6444
DB	2306	GGACTATAAAGATACCAGCGGTTTCCCTCTGGAAGCTCCCTCTGCGCTCTCTCTGTCCG	2247
QY	6445	ACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAGCGTGGGCTTCT	6504
DB	2246	ACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAGCGTGGGCTTCT	2187
QY	6505	CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCCAAGCTGGGCTGT	6564
DB	2186	CAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCCAAGCTGGGCTGT	2127
QY	6565	GTGCACGAACCCCGCTTCAGCGGACCGCTCGGCTTATCCGGTAACTATGCTCTTGA	6624
DB	2126	GTGCACGAACCCCGCTTCAGCGGACCGCTCGGCTTATCCGGTAACTATGCTCTTGA	2067
QY	6625	TCCAAACCCGGTAAGACACGACTTATCGCCACTGGGACAGCCACTGGTAACAGGATTAC	6684
DB	2066	TCCAAACCCGGTAAGACACGACTTATCGCCACTGGGACAGCCACTGGTAACAGGATTAC	2007
QY	6685	AGAGCAGGTATGTAGCGGCTACAGAGTCTTGAAGTGTGGCTTAACCTACGGCTAC	6744
DB	2006	AGAGCAGGTATGTAGCGGCTACAGAGTCTTGAAGTGTGGCTTAACCTACGGCTAC	1947
QY	6745	ACTAGAAGACACAGTATTTGGTATCTCGCTCTGCTGAGCCAGTACCTTCGGAAGAA	6804
DB	1946	ACTAGAAGACACAGTATTTGGTATCTCGCTCTGCTGAGCCAGTACCTTCGGAAGAA	1887
QY	6805	GTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCG	6864
DB	1886	GTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCG	1827
QY	6865	AGCAGCAGATTACCGCGCAAAAAGAGATCTCAAGAGATCCTTTGATCTTTTCTACG	6924
DB	1826	AGCAGCAGATTACCGCGCAAAAAGAGATCTCAAGAGATCCTTTGATCTTTTCTACG	1767
QY	6925	GGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTGGTCATGAGATTATCA	6984
DB	1766	GGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTGGTCATGAGATTATCA	1707
QY	6985	AAAAGATCTTCACTAGATCTCTTTTAAATTAAGTGTAAATCAATCTAAAGT	7044
DB	1706	AAAAGATCTTCACTAGATCTCTTTTAAATTAAGTGTAAATCAATCTAAAGT	1647
QY	7045	ATATATGAGTAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGCACCTATCTCA	7104
DB	1646	ATATATGAGTAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGCACCTATCTCA	1587
QY	7105	GGCATCTGCTATTTCTGTTTCATCTAGTGTGCTTCACTCCCGCTGCTGTAGATACTAG	7164
DB	1586	GGCATCTGCTATTTCTGTTTCATCTAGTGTGCTTCACTCCCGCTGCTGTAGATACTAG	1527
QY	7165	ATACGGAGGGCTTACCATCTGGCCCGCAGTGTCAATGATACCGCGAGACCCAGCTCA	7224
DB	1526	ATACGGAGGGCTTACCATCTGGCCCGCAGTGTCAATGATACCGCGAGACCCAGCTCA	1467
QY	7225	CGGGTCCAGATTTATCAGCAATAAACACGACGCGGAGGCGGAGCGCAGAGAGTGT	7284
DB	1466	CGGGTCCAGATTTATCAGCAATAAACACGACGCGGAGGCGGAGCGCAGAGAGTGT	1407
QY	7285	CTTGCAATCTTATCGGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGAGTAA	7344
DB	1406	CTTGCAATCTTATCGGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGAGTAA	1347
QY	7345	AGTTGCGCAGTTAATAGTTTGGCAACCGTTGTCATTGCTGAGGAGATCTGTGTGTC	7404
DB	1346	AGTTGCGCAGTTAATAGTTTGGCAACCGTTGTCATTGCTGAGGAGATCTGTGTGTC	1287

7405 CGCTCGTGGTGGTATGGCTTCATTCAGCTCCGTTCCCAAGATCAAGGCGAGTTACA 7464  
1286 CGCTCGTGGTGGTATGGCTTCATTCAGCTCCGTTCCCAAGATCAAGGCGAGTTACA 1227  
7465 TGATCCCGCATGTTGTCACAAAAGCGGTTAGCTCTCGGTCCTCGGATCGTTGTCAGA 7524  
1226 TGATCCCGCATGTTGTCACAAAAGCGGTTAGCTCTCGGTCCTCGGATCGTTGTCAGA 1167  
7525 AGTAAGTTGCGCGAGGTTATCACTCATGTTATGGCAGCACTGCATATTTCTTACT 7584  
1166 AGTAAGTTGCGCGAGGTTATCACTCATGTTATGGCAGCACTGCATATTTCTTACT 1107  
7585 GTCATGCCATCCGTAAGATGTTTCTGTGACCTGGTGAGTACTCAACCAAGTCATCTGA 7644  
1106 GTCATGCCATCCGTAAGATGTTTCTGTGACCTGGTGAGTACTCAACCAAGTCATCTGA 1047  
7645 GAATAGTGTATCGCGGACCGAGTTGCTCTTGCCCGCGTCAACAGCGGATATACCGCG 7704  
1046 GAATAGTGTATCGCGGACCGAGTTGCTCTTGCCCGCGTCAACAGCGGATATACCGCG 987  
7705 CCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAAAGTTCTTCGGGCGGAAACCTC 7764  
986 CCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAAAGTTCTTCGGGCGGAAACCTC 927  
7765 TCAGGATCTTACCGTGTGAGATCCAGTTCGATGTAAACCACTGTGCAACCAACTGA 7824  
926 TCAGGATCTTACCGTGTGAGATCCAGTTCGATGTAAACCACTGTGCAACCAACTGA 867  
7825 TCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAAACAGGAAGCGCAAAAT 7884  
866 TCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAAACAGGAAGCGCAAAAT 807  
7885 CCGCAAAAAGGAATAAGGCGACACGGAATGTTGAATATCTATCTTCTTCTTTT 7944  
806 CCGCAAAAAGGAATAAGGCGACACGGAATGTTGAATATCTATCTTCTTCTTTT 747  
7945 CAATATTATCAAGCATTTATCAGGTTTATGTTCTCATGCGGATACATATTGTAATGT 8004  
746 CAATATTATCAAGCATTTATCAGGTTTATGTTCTCATGCGGATACATATTGTAATGT 687  
8005 ATTTAGAAAAATAACAAATAGGCGTTCGCGGACATTTCCCGAAAGTGCACCTGAC 8064  
686 ATTTAGAAAAATAACAAATAGGCGTTCGCGGACATTTCCCGAAAGTGCACCTGAC 627  
8065 GTCTAAGAAACCATTTATCATGACATTAACCTTATAAAATAGCGGTATCAGAGGCC 8124  
626 GTCTAAGAAACCATTTATCATGACATTAACCTTATAAAATAGCGGTATCAGAGGCC 567  
8125 TTTCTCTTCAAGATTAATTTCTCATGTTTGCAGAGTTCATCGATAAGCT 8176  
566 TTTCTCTTCAAG----AATTTCTCATGTTTGCAGAGTTCATCGATAAGCT 519

RESULT 9

US-09-245-497-4/c

Sequence 4, Application US/09245497

Patent No. 6083750

GENERAL INFORMATION:

APPLICANT: Chamberlain, Jeffrey S.

APPLICANT: Amalitano, Andrea

APPLICANT: Hauser, Michael A.

APPLICANT: Kumar-Singh, Rajendra

APPLICANT: Hartigan-O'Connor, Dennis J.

TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States Of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/245,497  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/735,609  
APPLICATION NUMBER: <B> FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ingolia, Diane E.  
REGISTRATION NUMBER: 40,027  
REFERENCE/DOCKET NUMBER: UM-02484  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34303 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "DNA"  
US-09-245-497-4

Query Match 28.0%; Score 2310.8; DB 3; Length 34303;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2326; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 5845 CTGCTCCGCGGTTTCGGTGATGACGGTGAACAACTCTGACATGACAGTCCCGGAGAC 5904  
DB 2846 CTGCTCCGCGGTTTCGGTGATGACGGTGAACAACTCTGACATGACAGTCCCGGAGAC 2787  
QY 5905 GGTACACAGTTGTCTGTAAGCGGATGCGGAGCAGACCAAGCCGTCAGGCGCGTTCAGC 5964  
DB 2786 GGTACACAGTTGTCTGTAAGCGGATGCGGAGCAGACCAAGCCGTCAGGCGCGTTCAGC 2727  
QY 5965 GGGTGTTCGGCGGTCTCGGGGCGCAGCCATGACCCAGTCACGTAGCGGATGAGTGA 6024  
DB 2726 GGGTGTTCGGCGGTCTCGGGGCGCAGCCATGACCCAGTCACGTAGCGGATGAGTGA 2667  
QY 6025 TACTGGCTTAACATATGCGGATCAGAGCAGATTGTACTGAGAGTGCCACCATATCGGTTGT 6084  
DB 2666 TACTGGCTTAACATATGCGGATCAGAGCAGATTGTACTGAGAGTGCCACCATATCGGTTGT 2607  
QY 6085 GAAATACCGCACAGATGCGTAAGGAGAAAATAACCGCATCAGCGCGCTCTTCGCTTCCTCG 6144  
DB 2606 GAAATACCGCACAGATGCGTAAGGAGAAAATAACCGCATCAGCGCGCTCTTCGCTTCCTCG 2547  
QY 6145 CTCACTGACTCGCTCGGTCGTTTCGGTGGCGGAGCGGTTATCAGTCACTCAAG 6204  
DB 2546 CTCACTGACTCGCTCGGTCGTTTCGGTGGCGGAGCGGTTATCAGTCACTCAAG 2487  
QY 6205 GCGGTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAACAAATGTGAGCAAAA 6264  
DB 2486 GCGGTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAACAAATGTGAGCAAAA 2427  
QY 6265 GCGCAGCAAAAAGCCAGGAACCGTAAAAAGCGCGGTTGCTGGCGTTTTCATAGGCTC 6324  
DB 2426 GCGCAGCAAAAAGCCAGGAACCGTAAAAAGCGCGGTTGCTGGCGTTTTCATAGGCTC 2367  
QY 6325 CGCCCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGTGCAGAAACCCGACA 6384  
DB 2366 CGCCCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGTGCAGAAACCCGACA 2307  
QY 6385 GGACTATAAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTGTCGCTTCCTTTCG 6444  
DB 2306 GGACTATAAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTGTCGCTTCCTTTCG 2247  
QY 6445 ACCTGCGGCTTACCGGATACCTGTCCGCTTTCCTCCCTTCGGGAAGCGTGCCTTTCT 6504

Db 2246 ACCCTGCGGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGAAGCGTGGCGCTTCT 2187  
Qy 6505 CATAGCTCAGCGTGTAGGTATCTCAGTTCGCTGAGTGTGCTCGCTCAAGCTGGGCTGT 6564  
Db 2186 CAACTGCTCAGCGTGTAGGTATCTCAGTTCGCTGAGTGTGCTCGCTCAAGCTGGGCTGT 2127  
Qy 6565 GTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGCTTTGAG 6624  
Db 2126 GTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGCTTTGAG 2067  
Qy 6625 TCCAAACCGGTAAAGACAGCTTATCGCCACTGCGCAGCAGCACTGTGTAAACAGATTAGC 6684  
Db 2066 TCCAAACCGGTAAAGACAGCTTATCGCCACTGCGCAGCAGCACTGTGTAAACAGATTAGC 2007  
Qy 6685 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTTGAAGTGTGGCTTAACTACGGCTAC 6744  
Db 2006 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTTGAAGTGTGGCTTAACTACGGCTAC 1947  
Qy 6745 ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTGTACCTTCGGAAGAAGA 6804  
Db 1946 ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTGTACCTTCGGAAGAAGA 1887  
Qy 6805 GTTGTAGTCTTGTATCCGGCAAAACAAACACCGCTGCTGAGCGGTGTTTTTTTGTTCG 6864  
Db 1886 GTTGTAGTCTTGTATCCGGCAAAACAAACACCGCTGCTGAGCGGTGTTTTTTTGTTCG 1827  
Qy 6865 AAGCAGAGATTACCGCAGAAAAAAGGATCTCAAGAAGTCTCTTGTATCTTTCTACG 6924  
Db 1826 AAGCAGAGATTACCGCAGAAAAAAGGATCTCAAGAAGTCTCTTGTATCTTTCTACG 1767  
Qy 6925 GGTCTGACGCTCAGTGTGAAGCAAACTCAGTTAAGGATTTTGTTCATCAGATTATCA 6984  
Db 1766 GGTCTGACGCTCAGTGTGAAGCAAACTCAGTTAAGGATTTTGTTCATCAGATTATCA 1707  
Qy 6985 AAAAGGATCTTACCTAGATCTTTTAAATTAATAAGTCTTTTAAATCAATCTAAAGT 7044  
Db 1706 AAAAGGATCTTACCTAGATCTTTTAAATTAATAAGTCTTTTAAATCAATCTAAAGT 1647  
Qy 7045 ATATATGATTAACCTTGTGACAGTTACCAATGCTTAACTCAGTGGGACCTATCTCA 7104  
Db 1646 ATATATGATTAACCTTGTGACAGTTACCAATGCTTAACTCAGTGGGACCTATCTCA 1587  
Qy 7105 GCGATCTGTATTTGCTTATCCATAGTGTGCTGACTCCCGCTGCTGATAGTAACCTACG 7164  
Db 1586 GCGATCTGTATTTGCTTATCCATAGTGTGCTGACTCCCGCTGCTGATAGTAACCTACG 1527  
Qy 7165 ATACGGAGGGCTTACCTTGTGCGCCAGTGTGCAATGATACCGGAGACCCAGCTCA 7224  
Db 1526 ATACGGAGGGCTTACCTTGTGCGCCAGTGTGCAATGATACCGGAGACCCAGCTCA 1467  
Qy 7225 CCGGCTCCAGATTATCAGCAATAAACAGCAGCGGAGCGGCGGAGCGAGAGTGGT 7284  
Db 1466 CCGGCTCCAGATTATCAGCAATAAACAGCAGCGGAGCGGCGGAGCGAGAGTGGT 1407  
Qy 7285 CTGCAACTTATCCGCTCCATCAGTCTATTATTTGCTCGGAGCTAGAGTAAGT 7344  
Db 1406 CTGCAACTTATCCGCTCCATCAGTCTATTATTTGCTCGGAGCTAGAGTAAGT 1347  
Qy 7345 AGTTCGCGAGTTAATAGTTTGGCAACGTTGTGCAATGCTGAGGCACTGCTGTGTC 7404  
Db 1346 AGTTCGCGAGTTAATAGTTTGGCAACGTTGTGCAATGCTGAGGCACTGCTGTGTC 1287  
Qy 7405 CCGCTCGTGTGTTGGTATGCTTCACTGCTCCGCTTCCCAACCATCAAGCGGAGTTACA 7464  
Db 1286 CCGCTCGTGTGTTGGTATGCTTCACTGCTCCGCTTCCCAACCATCAAGCGGAGTTACA 1227  
Qy 7465 TGATCCCGCATGTTGTGCAAAAAAGCGGTAGTCTCTTCCGTCCTCCGATGTTGTGAGA 7524  
Db 1226 TGATCCCGCATGTTGTGCAAAAAAGCGGTAGTCTCTTCCGTCCTCCGATGTTGTGAGA 1167  
Qy 7525 AGTAAGTTGGCGGAGTGTATCACTCATGTTGTGCGAGCACTGCATAATCTCTTACT 7584

Db 1166 AGTAAGTTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTACT 1107  
Qy 7585 GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTGAGTACTCAACCAAGTCAATCTGA 7644  
Db 1106 GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTGAGTACTCAACCAAGTCAATCTGA 1047  
Qy 7645 GAATAGTGTATGGCGGACCGAGTGTCTCTTGTCCCGCGCTCAACACGGGATATACCGG 7704  
Db 1046 GAATAGTGTATGGCGGACCGAGTGTCTCTTGTCCCGCGCTCAACACGGGATATACCGG 987  
Qy 7705 CCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAACGTTCTTTCGGGGCGGAAACTC 7764  
Db 986 CCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAACGTTCTTTCGGGGCGGAAACTC 927  
Qy 7765 TCAAGGATCTTACCGCTGTGTGAGTCCAGTTCGATTAACCCACTGTCGACCCCACTGA 7824  
Db 926 TCAAGGATCTTACCGCTGTGTGAGTCCAGTTCGATTAACCCACTGTCGACCCCACTGA 867  
Qy 7825 TCTTCAGCATCTTTTACTTTTCCACGCTTTCTGGGTGAGCAAAAACAGAAAGGCAAAAT 7884  
Db 866 TCTTCAGCATCTTTTACTTTTCCACGCTTTCTGGGTGAGCAAAAACAGAAAGGCAAAAT 807  
Qy 7885 GCGCAAAAACAGGAAATAAGGCGCACCGAAATGTTGAATCTCATCTCTTCTCTTTT 7944  
Db 806 GCGCAAAAACAGGAAATAAGGCGCACCGAAATGTTGAATCTCATCTCTTCTCTTTT 747  
Qy 7945 CAATATATTGAGCATTTTATCAGGCTTATGCTCATGCGGTATACATATTTGAATCT 8004  
Db 746 CAATATATTGAGCATTTTATCAGGCTTATGCTCATGCGGTATACATATTTGAATCT 687  
Qy 8005 ATTTAGAAAAATAAACAATAGGCTTCCGCGCACATTTCCCGGAAAGTGCACCTGAC 8064  
Db 686 ATTTAGAAAAATAAACAATAGGCTTCCGCGCACATTTCCCGGAAAGTGCACCTGAC 627  
Qy 8065 GCTTAGAAACCAATTAATCATGACATTAACCTATAAAAATAGGCTATCAGAGGCC 8124  
Db 626 GCTTAGAAACCAATTAATCATGACATTAACCTATAAAAATAGGCTATCAGAGGCC 567  
Qy 8125 TTTTCGCTTCAAGAATAATTTCTCATGTTTGACAGCTTATCATCGATAAGCT 8176  
Db 566 TTTTCGCTTCAAG-----AATTCATGTTTGACAGCTTATCATCGATAAGCT 519

## RESULT 10

US-09-562-919-4/c

; Sequence 4, Application US/09562919

; Patent No. 6451596

; GENERAL INFORMATION:

; APPLICANT: Chamberlain, Jeffrey S.

; Amalfitano, Andrea

; Hauser, Michael A.

; Kumar-Singh, Rajendra

; Hartigan-O'Connor, Dennis J.

; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Medlen &amp; Carroll, LLP

; STREET: 220 Montgomery Street, Suite 2200

; CITY: San Francisco

; STATE: California

; COUNTRY: United States Of America

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/562,919

; FILING DATE: 02-May-2000

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/735,609





QY 7705 CCACATAGCAGAACTTTAAAGTGTCTATCTATTCGAAAAAGTCTTCGGGGCGAAAACTC 7764  
Db 986 CCACATAGCAGAACTTTAAAGTGTCTATCTATTCGAAAAAGTCTTCGGGGCGAAAACTC 927  
QY 7765 TCAAGGATCTTACCGCTGTTCAGATCCAGTTCGATGTAACCCACTCGTGACACCAACTGA 7824  
Db 926 TCAAGGATCTTACCGCTGTTCAGATCCAGTTCGATGTAACCCACTCGTGACACCAACTGA 867  
QY 7825 TCTTCAGCATCTTTTACCTTCCACAGCGCTTCTCGGTGAGCAAAAAACAGGAAGCAAAAT 7884  
Db 866 TCTTCAGCATCTTTTACCTTCCACAGCGCTTCTCGGTGAGCAAAAAACAGGAAGCAAAAT 807  
QY 7885 GCGCCAAAAAGGGAATAAGGCGGACACGGAATGTTGAATACATCATCTTCTCTCTTTT 7944  
Db 806 GCGCCAAAAAGGGAATAAGGCGGACACGGAATGTTGAATACATCATCTTCTCTCTTTT 747  
QY 7945 CAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAATGT 8004  
Db 746 CAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAATGT 687  
QY 8005 ATTTAGAAAAATAACAAATAGGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGAC 8064  
Db 686 ATTTAGAAAAATAACAAATAGGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGAC 627  
QY 8065 GTCTAAGAAACCAATTATTATCATGACATTAACCTATTAATAATAGGCGGTATCAGGAGGCC 8124  
Db 626 GTCTAAGAAACCAATTATTATCATGACATTAACCTATTAATAATAGGCGGTATCAGGAGGCC 567  
QY 8125 TTTGCTCTTAAGAAATTAATTTCTCATGTTTACAGCTTATCATGATAAGCT 8176  
Db 566 TTTGCTCTTAAG---AATTCTCATGTTTACAGCTTATCATGATAAGCT 519

## RESULT 11

US-08-125-462-2/c  
; Sequence 2, Application US/08125462  
; Patent No. 5840840  
; GENERAL INFORMATION:  
; APPLICANT: Rybak, Susanna M.  
; APPLICANT: Youle, Richard J.  
; APPLICANT: Newton, Dianne L.  
; APPLICANT: Nicholls, Peter J.  
; TITLE OF INVENTION: Selective RNase Cytotoxic Reagents  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESSES:  
; ADDRESS: Townsend and Townsend Khourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/125,462  
; FILING DATE: 22-SEP-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen L.  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 15280-110-3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6727 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1..6727  
; OTHER INFORMATION: /standard\_name="PET-11d-ANG-E6"  
; US-08-125-462-2

Query Match 27.9%; Score 2302; DB 2; Length 6727;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2328; Conservative 0; Mismatches 0; Indels 6; Gaps 2;

QY 5845 CTGCTTCGGCGCTTCGCTGATGACGGTCAAAACCTCTCACACATGCAGCTCCCGAGAC 5904  
Db 3821 CTGCTTCGGCGCTTCGCTGATGACGGTCAAAACCTCTCACACATGCAGCTCCCGAGAC 3762  
QY 5905 GGTCAACAGCTTGTCTAAGCGGATGCGGGAGCAGAACGCCGTCAGGGCGCGTCAGC 5964  
Db 3761 GGTCAACAGCTTGTCTAAGCGGATGCGGGAGCAGAACGCCGTCAGGGCGCGTCAGC 3702  
QY 5965 GGTGTTCGGCGCTTCGCTAAGCGGATGCGGGAGCAGAACGCCGTCAGGGCGCGTCAGC 6024  
Db 3701 GGTGTTCGGCGCTTCGCTAAGCGGATGCGGGAGCAGAACGCCGTCAGGGCGCGTCAGC 3642  
QY 6025 TACTGCTTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACC--ATATGCGGT 6082  
Db 3641 TACTGCTTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATATGCGGT 3582  
QY 6083 GTGAATAACGCACAGATCGGTAAAGAGAAATACCGCATCAGGCGCTCTTCGGCTTCCT 6142  
Db 3581 GTGAATAACGCACAGATCGGTAAAGAGAAATACCGCATCAGGCGCTCTTCGGCTTCCT 3522  
QY 6143 CGCTCACTGACTCGCTCGCTCGCTTCGGCTTCGGCGAGCGGTATCAGCTCACTCAA 6202  
Db 3521 CGCTCACTGACTCGCTCGCTCGCTTCGGCTTCGGCGAGCGGTATCAGCTCACTCAA 3462  
QY 6203 AGCGGTAAATACGGTTATCCACAGAAATAGGGGATAACCCAGAGAAAGAACATGTGAGCAA 6262  
Db 3461 AGCGGTAAATACGGTTATCCACAGAAATAGGGGATAACCCAGAGAAAGAACATGTGAGCAA 3402  
QY 6263 AAGGCCAGCAAAAGGCCAGGAAACCGTAAAGAGCGCGCTTCGCTGGCGTTTTTCCATAGGC 6322  
Db 3401 AAGGCCAGCAAAAGGCCAGGAAACCGTAAAGAGCGCGCTTCGCTGGCGTTTTTCCATAGGC 3342  
QY 6323 TCCGCCCTTCGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTTCGCGAAACCCGA 6382  
Db 3341 TCCGCCCTTCGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTTCGCGAAACCCGA 3282  
QY 6383 CAGGACTATAAGATACAGGCGTTTTCCCGCTGGAGCTCCCTCGTGGCTCTCCTGTTC 6442  
Db 3281 CAGGACTATAAGATACAGGCGTTTTCCCGCTGGAGCTCCCTCGTGGCTCTCCTGTTC 3222  
QY 6443 CGACCTTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTT 6502  
Db 3221 CGACCTTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTT 3162  
QY 6503 CTCTAGCTCAGCTCTAGGTATCTCAGTTCCGTTAGGTGCTGCTCCAGCTGGGCT 6562  
Db 3161 CTCTAGCTCAGCTCTAGGTATCTCAGTTCCGTTAGGTGCTGCTCCAGCTGGGCT 3102  
QY 6563 GTGTGACGAAACCCCGCTTCAGCGCGAGCGCTGCGCTTATCCGGTAACTATCTGCTTCG 6622  
Db 3101 GTGTGACGAAACCCCGCTTCAGCGCGAGCGCTGCGCTTATCCGGTAACTATCTGCTTCG 3042  
QY 6623 AGTCCAAACCGGTAAAGACACGACTTATCCCACTGCGACAGCCACTGTTAAAGATTA 6682  
Db 3041 AGTCCAAACCGGTAAAGACACGACTTATCCCACTGCGACAGCCACTGTTAAAGATTA 2982  
QY 6683 GCAGAGGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAATAGGCT 6742  
Db 2981 GCAGAGGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAATAGGCT 2922

6743 ACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAA 6802  
Db  
2921 ACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAA 2862  
QY  
6803 GAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCACCGCTGCTAGCGGTGGTTTTTTTGT 6862  
Db  
2861 GAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCACCGCTGCTAGCGGTGGTTTTTTGT 2802  
QY  
6863 GCAAGCAGCAGATTACGGCGAGAAAAAGAGATCTCAAGAGATCTCTTTCATCTTTCTA 6922  
Db  
2801 GCAAGCAGCAGATTACGGCGAGAAAAAGAGATCTCAAGAGATCTCTTTCATCTTTCTA 2742  
QY  
6923 CGGGCTCTGAGCTCAGTGAAGCAAAACCTCAGTTAAGGATTTTGGTCTCATGAGATTAT 6982  
Db  
2741 CGGGCTCTGAGCTCAGTGAAGCAAAACCTCAGTTAAGGATTTTGGTCTCATGAGATTAT 2682  
QY  
6983 CAAAAGGATCTTCACCTAGATCTCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAA 7042  
Db  
2681 CAAAAGGATCTTCACCTAGATCTCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAA 2622  
QY  
7043 GTATATAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGACCTATCT 7102  
Db  
2621 GTATATAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGACCTATCT 2562  
QY  
7103 CAGGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTA 7162  
Db  
2561 CAGGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTA 2502  
QY  
7163 CGATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGCCACGCT 7222  
Db  
2501 CGATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGCCACGCT 2442  
QY  
7223 CACCGCTCCAGATTATACCAATTAACACAGCCAGCGGAGGCGGAGCGAGAGTG 7282  
Db  
2441 CACCGCTCCAGATTATACCAATTAACACAGCCAGCGGAGGCGGAGCGAGAGTG 2382  
QY  
7283 GTCCTGCAACTTATTCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAA 7342  
Db  
2381 GTCCTGCAACTTATTCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAA 2322  
QY  
7343 GTAGTTCCGCAATTAATAGTTTGGCAACAGTGTTCATTTGTCGAGGAGCATCGTGTGT 7402  
Db  
2321 GTAGTTCCGCAATTAATAGTTTGGCAACAGTGTTCATTTGTCGAGGAGCATCGTGTGT 2262  
QY  
7403 CACGCTCGTCTGTTGGTAGTGGCTTCATTCAGCTCCGCTTCCACAGATCAAGCGAGTTA 7462  
Db  
2261 CACGCTCGTCTGTTGGTAGTGGCTTCATTCAGCTCCGCTTCCACAGATCAAGCGAGTTA 2202  
QY  
7463 CATGATCCCGCATGTTGTCGAAAAGCGGTAGCTCTTCGTCCTCCGATCGTTGTC 7522  
Db  
2201 CATGATCCCGCATGTTGTCGAAAAGCGGTAGCTCTTCGTCCTCCGATCGTTGTC 2142  
QY  
7523 GAAGTAAGTTGCGCCAGCTTTATCACTCATGTTATGGCAGCACTGCATAATTTCTTTA 7582  
Db  
2141 GAAGTAAGTTGCGCCAGCTTTATCACTCATGTTATGGCAGCACTGCATAATTTCTTTA 2082  
QY  
7583 CTGTCATGCCATCCGTAGATGCTTTTCTGTCAGCTGTTGAGTACTCAACCAAGTCATCT 7642  
Db  
2081 CTGTCATGCCATCCGTAGATGCTTTTCTGTCAGTGTGATCTCAACCAAGTCATCT 2022  
QY  
7643 GAGATAGTGTATGGGCGACCGAGTTGCTCTTCGCGCGGTCAACACGGGATAATACCG 7702  
Db  
2021 GAGATAGTGTATGGGCGACCGAGTTGCTCTTCGCGCGGTCAACACGGGATAATACCG 1962  
QY  
7703 CGCCCATAGCAGAACTTTAAAGTGCTCATCTATGGAAAAAGTTCTTCGGGGGAAAAAC 7762  
Db  
1961 CGCCCATAGCAGAACTTTAAAGTGCTCATCTATGGAAAAAGTTCTTCGGGGGAAAAAC 1902  
QY  
7763 TCTCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGTGACCCCACT 7822  
Db  
1901 TCTCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGTGACCCCACT 1842  
QY  
7823 GATCTTCAGATCTTTTACTTTCACCAAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAA 7882

1841 GATCTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAA 1782  
QY  
7883 ATGCGCGAAAAAAGGAATAGGGCGACACGGAATGTTGAATACTCATACTCTTCCCTTT 7942  
Db  
1781 ATGCGCGAAAAAAGGAATAGGGCGACACGGAATGTTGAATACTCATACTCTTCCCTTT 1722  
QY  
7943 TTCAATATTTTGAAGCACTTTTATCAGGGTTATTTCTCATGAGCGGATACATATTTGAAT 8002  
Db  
1721 TTCAATATTTTGAAGCACTTTTATCAGGGTTATTTCTCATGAGCGGATACATATTTGAAT 1662  
QY  
8003 GTATTTAGAAAAATAACAAATAGGGTTCCGCGACATTTCCCGGAAAAAGTGCACCTG 8062  
Db  
1661 GTATTTAGAAAAATAACAAATAGGGTTCCGCGACATTTCCCGGAAAAAGTGCACCTG 1602  
QY  
8063 AGCTCTAGAAAAACCAATTTATTCATGACATTAACCTATAAAATAGGCGTATCACGAGGC 8122  
Db  
1601 AGCTCTAGAAAAACCAATTTATTCATGACATTAACCTATAAAATAGGCGTATCACGAGGC 1542  
QY  
8123 CTTTTCTGCTTTCAAGAAATTAATTTCTCATGTTTGAAGCTTATCATTCGATAAGCT 8176  
Db  
1541 CTTTTCTGCTTTCAAG---AATTTCTGTTTGAAGCTTATCATTCGATAAGCT 1492

## RESULT 12

US-08-891-848-2/c  
; Sequence 2, Application US/08891848  
; Patent No. 5955073  
; GENERAL INFORMATION:  
; APPLICANT: Rybak, Susanna M.  
; APPLICANT: Youle, Richard J.  
; APPLICANT: Newton, Dianne L.  
; APPLICANT: Nicholls, Peter J.  
; TITLE OF INVENTION: Selective Rhase Cytotoxic Reagents  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/891.848  
; FILING DATE: No. 5955073 yet assigned  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/125,462  
; FILING DATE: 22-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/014,082  
; FILING DATE: 04-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/779,195  
; FILING DATE: 22-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/510,696  
; FILING DATE: 20-APR-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 015280-110310US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6727 base pairs



QY 7823 GATCTTCAGCATCTTTTACTTTTACACGCGTTTCTGGTGAGCAAAACAGGAAGGCAAA 7882  
Db 1841 GATCTTCAGCATCTTTTACTTTTACACGCGTTTCTGGTGAGCAAAACAGGAAGGCAAA 1782  
QY 7883 ATGCCGCAAAAGGGAATTAAGGCGGACAGGAAATGTTGAATCTCTTCTCTTT 7942  
Db 1781 ATGCCGCAAAAGGGAATTAAGGCGGACAGGAAATGTTGAATCTCTTCTCTTT 1722  
QY 7943 TTCAATATTATTAAGCATTTATCAGAGGTTATTTCTCATGAGCGGATACATATTTGAAT 8002  
Db 1721 TTCAATATTATTAAGCATTTATCAGAGGTTATTTCTCATGAGCGGATACATATTTGAAT 1662  
QY 8003 GTATTTAGAAAATAAACAATAGGGTTCCGGCGACATTTCCCGCAAAAGTGCACCTG 8062  
Db 1661 GTATTTAGAAAATAAACAATAGGGTTCCGGCGACATTTCCCGCAAAAGTGCACCTG 1602  
QY 8063 ACGTCTAAGAAACCATTTATCATGACATTAACCTATTAATAATAGGCGTATCACGAGGC 8122  
Db 1601 ACGTCTAAGAAACCATTTATCATGACATTAACCTATTAATAATAGGCGTATCACGAGGC 1542  
QY 8123 CCTTTCGCTTTCAAGAAATTAATCTCATGTTTGAAGCTTTATCATCGATTAAGCT 8176  
Db 1541 CCTTTCGCTTTCAAG---AATTTCTCATGTTTGAAGCTTTATCATCGATTAAGCT 1492

## RESULT 13

US-08-125-462-5/c  
; Sequence 5, Application US/08125462  
; Patent No. 5840840  
; GENERAL INFORMATION:  
; APPLICANT: Rybak, Susanna M.  
; APPLICANT: Koule, Richard J.  
; APPLICANT: Newton, Dianne L.  
; APPLICANT: Nicholls, Peter J.  
; TITLE OF INVENTION: Selective RNase Cytotoxic Reagents  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 22-SEP-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen L.  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 15280-110-3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6799 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEetical: NO  
; NAME/KEY: misc feature  
; LOCATION: 1..6799  
; OTHER INFORMATION: /standard\_name="pET-11d-E6-FB-EDN"  
US-08-125-462-5

Query Match 27.9%; Score 2302; DB 2; Length 6799;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2328; Conservative 0; Mismatches 0; Indels 6; Gaps 2;  
QY 5845 CTGCCCTCCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGAGCTCCCGGAGAC 5904  
Db 3893 CTGCCCTCCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGAGCTCCCGGAGAC 3834  
QY 5905 GGTACACAGCTTTGTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGCGCGTCA 5964  
Db 3833 GGTACACAGCTTTGTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGCGCGTCA 3774  
QY 5965 GGGTGTTCGGCGGTTTCGGGCGGAGCCATGACCCAGTCAAGTACCGATAGCGAGGTGA 6024  
Db 3773 GGGTGTTCGGCGGTTTCGGGCGGAGCCATGACCCAGTCAAGTACCGATAGCGAGGTGA 3714  
QY 6025 TACTGGCTTAATATGCGGCATCAGAGCAGATTGTCTAGAGTGCACC--ATAATGGCGT 6082  
Db 3713 TACTGGCTTAATATGCGGCATCAGAGCAGATTGTCTAGAGTGCACCATATATGCGGT 3654  
QY 6083 GTGAAATACCGCAGATGCGTAAGGAAATACCGCATCAGCGCTCTTCCGCTTCCT 6142  
Db 3653 GTGAAATACCGCAGATGCGTAAGGAAATACCGCATCAGCGCTCTTCCGCTTCCT 3594  
QY 6143 CGCTCAGTACCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 6202  
Db 3593 CGCTCAGTACCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 3534  
QY 6203 AGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAAGAACATGTAGCAA 6262  
Db 3533 AGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAAGAACATGTAGCAA 3474  
QY 6263 RAGGCCAGCAAAAGCCAGGAAACCGTAAAGGCGGTTCTGCTGCGCTTTTCCATAGGC 6322  
Db 3473 AAGGCCAGCAAAAGCCAGGAAACCGTAAAGGCGGTTCTGCTGCGCTTTTCCATAGGC 3414  
QY 6323 TCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGA 6382  
Db 3413 TCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGA 3354  
QY 6383 CAGGACTATAAGATACAGGCGTTTCCCTCGAAGCTCCCTCGCTCGCTCTCTGTC 6442  
Db 3353 CAGGACTATAAGATACAGGCGTTTCCCTCGAAGCTCCCTCGCTCGCTCTCTGTC 3294  
QY 6443 CGACCTCGCGCTTACCGGATACCTGTCGCTTTCTCCCTTCGGGAAGCGTGGCGTTT 6502  
Db 3293 CGACCTCGCGCTTACCGGATACCTGTCGCTTTCTCCCTTCGGGAAGCGTGGCGTTT 3234  
QY 6503 CTCATAGCTCAGCGTGTAGGTATCTCAGTCCGTTAGGTCTGCTCCAGCTGGCT 6562  
Db 3233 CTCATAGCTCAGCGTGTAGGTATCTCAGTCCGTTAGGTCTGCTCCAGCTGGCT 3174  
QY 6563 GTGTGACGAACCCCGTTACCGGATACCTGTCGCTTTCTCCCTTCGGGAAGCGTGGCGTTG 6622  
Db 3173 GTGTGACGAACCCCGTTACCGGATACCTGTCGCTTTCTCCCTTCGGGAAGCGTGGCGTTG 3114  
QY 6623 AGTCCAAACCCCGTAAAGACACGACTTATCGCACTGGCAGAGCCACTGTAACAGGATTA 6682  
Db 3113 AGTCCAAACCCCGTAAAGACACGACTTATCGCACTGGCAGAGCCACTGTAACAGGATTA 3054  
QY 6683 GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACAGGCT 6742  
Db 3053 GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACAGGCT 2994  
QY 6743 AACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGCTTACCTCCGAAAA 6802  
Db 2993 AACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGCTTACCTCCGAAAA 2934  
QY 6803 GAGTTGGTAGCTTTGATCCGCAAAAGAACACCGCTGGTAGCGGTGTTTTTTTGT 6862  
Db 2933 GAGTTGGTAGCTTTGATCCGCAAAAGAACACCGCTGGTAGCGGTGTTTTTTTGT 2874

QY 6863 GCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTCTA 6922  
Db 2873 GCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTCTA 2814  
QY 6923 CGGGGTCTGAGCCTCAGTGGACGAAACTCAGCTTAAGGATTTTGGTCATGAGATTAT 6982  
Db 2813 CGGGGTCTGAGCCTCAGTGGACGAAACTCAGCTTAAGGATTTTGGTCATGAGATTAT 2754  
QY 6983 CAAAAAGATCTTCAACCTAGATCCTTTTAAATTAATAAGTCTTTTAAATCAATCTAAA 7042  
Db 2753 CAAAAAGATCTTCAACCTAGATCCTTTTAAATTAATAAGTCTTTTAAATCAATCTAAA 2694  
QY 7043 GTATATATGAGTAACCTTGGTCTGACAGATTACCAATGCTTAATCAGTGAGGACCTATCT 7102  
Db 2693 GTATATATGAGTAACCTTGGTCTGACAGATTACCAATGCTTAATCAGTGAGGACCTATCT 2634  
QY 7103 CAGGATCTGCTATTTGCTTATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTA 7162  
Db 2633 CAGGATCTGCTATTTGCTTATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTA 2574  
QY 7163 CGATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATACCGGAGACCCAGCT 7222  
Db 2573 CGATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATACCGGAGACCCAGCT 2514  
QY 7223 CACGGCTCCAGATTATCAGCAATAACACGAGCGGAGGCGGAGGCGAGAGTG 7282  
Db 2513 CACGGCTCCAGATTATCAGCAATAACACGAGCGGAGGCGGAGGCGAGAGTG 2454  
QY 7283 GTCTGCAACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAAGCTAGATAA 7342  
Db 2453 GTCTGCAACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAAGCTAGATAA 2394  
QY 7343 GTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTCGATGCTGCGGATCGTGTGT 7402  
Db 2393 GTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTCGATGCTGCGGATCGTGTGT 2334  
QY 7403 CACGCTCGCTTTGGTATGCTTCAATCAGCTCCGTTCCCAACGATCAAGGCGAGTTA 7462  
Db 2333 CACGCTCGCTTTGGTATGCTTCAATCAGCTCCGTTCCCAACGATCAAGGCGAGTTA 2274  
QY 7463 CATGATCCCCATGTTGCAAAAAAGCGTTAGCTCCTTCGGTCTCCGATGTTGCA 7522  
Db 2273 CATGATCCCCATGTTGCAAAAAAGCGTTAGCTCCTTCGGTCTCCGATGTTGCA 2214  
QY 7523 GAAGTAAGTTGGCGCAGTGTATCAGTCATGTTATGGAGCAGCTCATTAATTTCTTTA 7582  
Db 2213 GAAGTAAGTTGGCGCAGTGTATCAGTCATGTTATGGAGCAGCTCATTAATTTCTTTA 2154  
QY 7583 CTGTCATGCCATCCGTAAGATGCTTTTCTGACTGTGTAGTACTCAACCAAGTCAATCT 7642  
Db 2153 CTGTCATGCCATCCGTAAGATGCTTTTCTGACTGTGTAGTACTCAACCAAGTCAATCT 2094  
QY 7643 GAGATAGTGTATGCGGCGACGAGTTCTCTTCCGCGGCTGACACGAGGATATACCG 7702  
Db 2093 GAGATAGTGTATGCGGCGACGAGTTCTCTTCCGCGGCTGACACGAGGATATACCG 2034  
QY 7703 CGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGAAGAAAGTCTTCCGGGCGAAAC 7762  
Db 2033 CGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGAAGAAAGTCTTCCGGGCGAAAC 1974  
QY 7763 TCTCAAGATCTTACCGTGTGAGATCCAGTTCGATTAACCCACTGTCGCAACCCAACT 7822  
Db 1973 TCTCAAGATCTTACCGTGTGAGATCCAGTTCGATTAACCCACTGTCGCAACCCAACT 1914  
QY 7823 GATCTTCAGCATCTTTTACTTTCACACGCTTCTGGGTGAGCAAAAACAGGAGGCAAA 7882  
Db 1913 GATCTTCAGCATCTTTTACTTTCACACGCTTCTGGGTGAGCAAAAACAGGAGGCAAA 1854  
QY 7883 ATGCCGCAAAAAGGAATAAGGGCGACACGGAATGTTGAATACTCATCTTCTCTTT 7942  
Db 1853 ATGCCGCAAAAAGGAATAAGGGCGACACGGAATGTTGAATACTCATCTTCTCTTT 1794  
QY 7943 TTCAATATATTGAAGCAATTTATCAGGCTTATTTGCTCATGAGCGGATACATATTGAAT 8002

## RESULT 14

US-08-891-848-5/c

; Sequence 5, Application US/08891848

; Patent No. 5955073

; GENERAL INFORMATION:

; APPLICANT: Rybak, Susanna M.

; APPLICANT: Youle, Richard J.

; APPLICANT: Newton, Dianne L.

; APPLICANT: Nicholls, Peter J.

; TITLE OF INVENTION: Selective RNase Cytotoxic Reagents

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/891,848

; FILING DATE: No. 5955073 yet assigned

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/125,462

; FILING DATE: 22-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/014,082

; FILING DATE: 04-FEB-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/779,195

; FILING DATE: 22-OCT-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/510,696

; FILING DATE: 20-APR-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Weber, Ellen Lauver

; REGISTRATION NUMBER: 32,762

; REFERENCE/DOCKET NUMBER: 015280-110310US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6799 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; FEATURE:

; NAME/KEY:

; LOCATION: 1..6799

; OTHER INFORMATION: /note= "pET-11d-E6-FB-EDN fusion

Query Match 27.9%; Score 2302; DB 2; Length 6799;  
Best Local Similarity 99.7%; Pred.No. 0;  
Matches 2328; Conservative 0; Mismatches 0; Indels 6

2933	Qy	Db	GAGTTGGTAGCTCTTGATCGGCGAAACAAACACCGCTCGTAGCGGTGGTTTTTTTGGTTT	2877
6863	Qy	Qy	GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTTA	6922
2873	Db	Db	GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTTA	2814
6923	Qy	Qy	CGGGTCTGACGCTCAGTGGAAACGAAATCAGCTTTAAAGGATTTTGGTCAUCAGATTAT	6982
2813	Db	Db	CGGGTCTGACGCTCAGTGGAAACGAAATCAGCTTTAAAGGATTTTGGTCAUCAGATTAT	2754
6983	Qy	Qy	CAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAA	7042
2753	Db	Db	CAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAA	2694
7043	Qy	Qy	GTATATAGTAAACTTGGTCTGCACAGTTACCAATGCTTAAATCAGTAGGCCACTTATCT	7102
2693	Db	Db	GTATATAGTAAACTTGGTCTGCACAGTTACCAATGCTTAAATCAGTAGGCCACTTATCT	2634
7103	Qy	Qy	CAGCGATCTGTCTATTTCGTTTCATCATAGTTGCCTGATCTCCCGTCGTGTAGATACTA	7162
2633	Db	Db	CAGCGATCTGTCTATTTCGTTTCATCATAGTTGCCTGATCTCCCGTCGTGTAGATACTA	2574
7163	Qy	Qy	CGATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGCGAGACCCACGCT	7222
2573	Db	Db	CGATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGCGAGACCCACGCT	2514
7223	Qy	Qy	CACCGCTCCAGATTATCAGCAATTAACACAGCCCGGAGGGCGCAGCGCAGAGTG	7282
2513	Db	Db	CACCGCTCCAGATTATCAGCAATTAACACAGCCCGGAGGGCGCAGCGCAGAGTG	2454
7283	Qy	Qy	GTCTCTCAACTTATCCGCTCCATCCAGTCTAATTAATTGTTGCCGGGAAGCTAGAGTAA	7342
2453	Db	Db	GTCTCTCAACTTATCCGCTCCATCCAGTCTAATTAATTGTTGCCGGGAAGCTAGAGTAA	2394
7343	Qy	Qy	GTAGTTGGCAGTTAATAGTTTGGCAAGTTTGTGTCATTTGCTGCAGGCATCGTGGTGT	7402
2393	Db	Db	GTAGTTGGCAGTTAATAGTTTGGCAAGTTTGTGTCATTTGCTGCAGGCATCGTGGTGT	2334
7403	Qy	Qy	CAGCTCTGCTGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGCGCAGTTA	7462
2333	Db	Db	CAGCTCTGCTGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGCGCAGTTA	2274
7463	Qy	Qy	CATGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCCTTCGTCCTCGATCGTTGTCA	7522
2273	Db	Db	CATGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCCTTCGTCCTCGATCGTTGTCA	2214
7523	Qy	Qy	GAGTAAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTCCTTA	7582
2213	Db	Db	GAGTAAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTCCTTA	2154
7583	Qy	Qy	CTGTCAATGCCATCCGTAAGATGCTTTTCTGTGATCGGTGAGTACTCAACCAAGTCAATCT	7642
2153	Db	Db	CTGTCAATGCCATCCGTAAGATGCTTTTCTGTGATCGGTGAGTACTCAACCAAGTCAATCT	2094
7643	Qy	Qy	GAGAAAGTGTATGCGGCGACCGAGTTGCTTCTTGCCCGCGCTCAACACGGGATAATACCG	7702
2093	Db	Db	GAGAAAGTGTATGCGGCGACCGAGTTGCTTCTTGCCCGCGCTCAACACGGGATAATACCG	2034
7703	Qy	Qy	CGCCAATAGCAGAACTTTAAAGTGCCTCATCTTGAAAAAGCTTCTTCGGGCGGAAAC	7762
2033	Db	Db	CGCCAATAGCAGAACTTTAAAGTGCCTCATCTTGAAAAAGCTTCTTCGGGCGGAAAC	1974
7763	Qy	Qy	TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCAACCACT	7822
1973	Db	Db	TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCAACCACT	1914
7823	Qy	Qy	GATCTTCAGATCTTTTACTTTCACAGCGTTTCTCGGTGAGCAAAAAACGAGAGGAAA	7882
1913	Db	Db	GATCTTCAGATCTTTTACTTTCACAGCGTTTCTCGGTGAGCAAAAAACGAGAGGAAA	1854
7883	Qy	Qy	ATGCCCAAAAAAGGAAATAAGGGCGACACGGAAATGTTGAATACATCATCTCTTCCTTT	7942
1853	Db	Db	ATGCCCAAAAAAGGAAATAAGGGCGACACGGAAATGTTGAATACATCATCTCTTCCTTT	1794



QY 7943 TTCAATATTATTGAAGCAATTTATCAGGTTATTGTCTCATGAGCGGATACATATTTGAAT 8002  
DB 1793 TTCAATATTATTGAAGCAATTTATCAGGTTATTGTCTCATGAGCGGATACATATTTGAAT 1734  
QY 8003 GTATTTAGAAAAATAAACAATAGGGTTCGCGCAATTTCCCGAAAAAGTGCCACCTG 8062  
DB 1733 GTATTTAGAAAAATAAACAATAGGGTTCGCGCAATTTCCCGAAAAAGTGCCACCTG 1674  
QY 8063 ACGTCTAGAAACCATTTATTCATGATTAACCTATAAAATAGCGGTATCAGAGGC 8122  
DB 1673 ACGTCTAGAAACCATTTATTCATGATTAACCTATAAAATAGCGGTATCAGAGGC 1614  
QY 8123 CCTTTGCTTTCAAGAATAATTTCTCATGTTTGTAGCAGCTTATCATCGATAAGCT 8176  
DB 1613 CCTTTGCTTTCAAG-----AATTCATGTTTGTAGCAGCTTATCATCGATAAGCT 1564

RESULT 15

US-08-757-439-1  
; Sequence 1, Application US/08757439  
; Patent No. 5866371  
; GENERAL INFORMATION:  
; APPLICANT: BADZIONG, Werner  
; APPLICANT: HABERMANN, Paul  
; APPLICANT: MOELLER, Joerg  
; APPLICANT: ARETZ, Werner  
; TITLE OF INVENTION: PROMOTER SYSTEM FOR THE PRODUCTION OF HETEROLOGOUS  
; TITLE OF INVENTION: PROTEINS IN HIGH YIELDS  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/757,439  
; FILING DATE: 27-NOV-1996  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 19544233.4  
; FILING DATE: 28-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SANDERCOCK, Colin G.  
; REGISTRATION NUMBER: 31,298  
; REFERENCE/DOCKET NUMBER: 18748/303/HOCE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8491 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-757-439-1

Query Match 27.9%; Score 2302; DB 2; Length 8491;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5845 CTGCCTCGCGCTTTTCGGTGATGACGGTGAACACCTCTGACACATGACAGCTCCCGGAGAC 5904  
DB 1508 CTGCCTCGCGCTTTTCGGTGATGACGGTGAACACCTCTGACACATGACAGCTCCCGGAGAC 1567

QY 5905 GGTCAAGCTTTGTCTTAAGCGGATCCCGGAGCAGACAAGCCCGTCAAGCGCGCTCAGC 5964  
DB 1568 GGTCAAGCTTTGTCTTAAGCGGATCCCGGAGCAGACAAGCCCGTCAAGCGCGCTCAGC 1627  
QY 5965 GGGTGTTCGGGGTGTTCGGGGCGAGCCATGACCCAGTCACTAGCGATAGCGAGTGA 5024  
DB 1628 GGGTGTTCGGGGTGTTCGGGGCGAGCCATGACCCAGTCACTAGCGATAGCGAGTGA 1687  
QY 6025 TACTGCTTTAACTATGCGGCATTCAGAGCAGATTGTACTGAGAGTGCACCATATCGGGTGT 6084  
DB 1688 TACTGCTTTAACTATGCGGCATTCAGAGCAGATTGTACTGAGAGTGCACCATATCGGGTGT 1747  
QY 6085 GAAATACCCGACAGATGCGTAAGAGGAAATACCGCATCAGGCGCTCTTCCGCTTCTCTCG 6144  
DB 1748 GAAATACCCGACAGATGCGTAAGAGGAAATACCGCATCAGGCGCTCTTCCGCTTCTCTCG 1807  
QY 6145 CTCACCTGACTCGCTCGCTCGGTTCGCTGCGCGAGCGGTATCAGCTCACTCAAAAG 6204  
DB 1808 CTCACCTGACTCGCTCGCTCGGTTCGCTGCGCGAGCGGTATCAGCTCACTCAAAAG 1867  
QY 6205 GCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAACATGTGAGCAAAA 6264  
DB 1868 GCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAACATGTGAGCAAAA 1927  
QY 6265 GGCACGAAAGCGCAGGAAACGTAAGAGCGGGTTCGCTGCGCTTTTCCATAGGCTC 6324  
DB 1928 GGCACGAAAGCGCAGGAAACGTAAGAGCGGGTTCGCTGCGCTTTTCCATAGGCTC 1987  
QY 6325 CGCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCGGACA 6384  
DB 1988 CGCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCGGACA 2047  
QY 6385 GGAATAAAGATACAGGCGTTTCCCTCGAGCTCCCTGCGGCTCTCTCTGTTCCG 6444  
DB 2048 GGAATAAAGATACAGGCGTTTCCCTCGAGCTCCCTGCGGCTCTCTCTGTTCCG 2107  
QY 6445 ACCCTGCCGTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTTCT 6504  
DB 2108 ACCCTGCCGTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTTCT 2167  
QY 6505 CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGTCGCTCCAGCTGGGCTGT 6564  
DB 2168 CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGTCGCTCCAGCTGGGCTGT 2227  
QY 6565 GTGACGAAACCCCGCTTCAGCCGCTTCAGCCGCTTATCCGTAACCTATCGTCTTGAG 6624  
DB 2228 GTGACGAAACCCCGCTTCAGCCGCTTATCCGTAACCTATCGTCTTGAG 2287  
QY 6625 TCCAAACCCGTAAGACACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTAGC 6684  
DB 2288 TCCAAACCCGTAAGACACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTAGC 2347  
QY 6685 AGAGCGAGGTATGAGCGGTGCTACAGAGTTCCTGAAGTGTGGCTTAACCTACGCTAC 6744  
DB 2348 AGAGCGAGGTATGAGCGGTGCTACAGAGTTCCTGAAGTGTGGCTTAACCTACGCTAC 2407  
QY 6745 ACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGA 6804  
DB 2408 ACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGA 2467  
QY 6805 GTTGGTAGCTTTGATCCGGCAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCG 6864  
DB 2468 GTTGGTAGCTTTGATCCGGCAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCG 2527  
QY 6865 AGCAGCAGATTACGCGCAAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACG 6924  
DB 2528 AGCAGCAGATTACGCGCAAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACG 2587  
QY 6925 GGGTCTGACGCTCAGTGGAAACGAAAACTCAGGTAAGGGATTGTCATGAGATTATCA 6984  
DB 2588 GGGTCTGACGCTCAGTGGAAACGAAAACTCAGGTAAGGGATTGTCATGAGATTATCA 2647

6985 AAAAGGATCTTACCTAGATCCCTTTTAAATATAAATGAAGTTTAAATCAATCTAAAGT 7044  
 2648 AAAAGGATCTTACCTAGATCCCTTTTAAATATAAATGAAGTTTAAATCAATCTAAAGT 2707  
 7045 ATATATGAGTAACCTTGGTGTGACAGTTACCAATGCTTAATCAGTGAAGGACCACTTCTCA 7104  
 2708 ATATATGAGTAACCTTGGTGTGACAGTTACCAATGCTTAATCAGTGAAGGACCACTTCTCA 2767  
 7105 GCGATCTGTCTATTTCTGTTTCATCCATAGTTGCTGCTGACTCCCGTCTGTTAGATAAATACG 7164  
 2768 GCGATCTGTCTATTTCTGTTTCATCCATAGTTGCTGCTGACTCCCGTCTGTTAGATAAATACG 2827  
 7165 ATACGGAGGGCTTACCATCTGGCCCGCAGTGTGCAATGATACCGGAGACCCACGCTCA 7224  
 2828 ATACGGAGGGCTTACCATCTGGCCCGCAGTGTGCAATGATACCGGAGACCCACGCTCA 2887  
 7225 CCGCTCCAGATTTATCAGCAATAAACCAGCCGCGAGGCGCGAGCGCAGAAAGTGGT 7284  
 2888 CCGCTCCAGATTTATCAGCAATAAACCAGCCGCGAGGCGCGAGCGCAGAAAGTGGT 2947  
 7285 CCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCCGGGAAGCTAGAGTAAGT 7344  
 2948 CCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCCGGGAAGCTAGAGTAAGT 3007  
 7345 AGTTCCGCGAGTTAATAGTTTGGCAAGCTTGTGCGCATGCTGCAGCATCGTGGTGTCA 7404  
 3008 AGTTCCGCGAGTTAATAGTTTGGCAAGCTTGTGCGCATGCTGCAGCATCGTGGTGTCA 3067  
 7405 CGCTCGTGTGTTGATGGCTTTCATTCAGCTCCGGTCCCAAGCATCAAGCGGAGTTACA 7464  
 3068 CGCTCGTGTGTTGATGGCTTTCATTCAGCTCCGGTCCCAAGCATCAAGCGGAGTTACA 3127  
 7465 TGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTTCCGTCCTCGATCGTGTGTCAGA 7524  
 3128 TGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTTCCGTCCTCGATCGTGTGTCAGA 3187  
 7525 AGTAAGTTGGCGCAGTGTTCATCTCATGTTATGCGCAGCTGCAATAATCTCTTACT 7584  
 3188 AGTAAGTTGGCGCAGTGTTCATCTCATGTTATGCGCAGCTGCAATAATCTCTTACT 3247  
 7585 GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTGA 7644  
 3248 GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTGA 3307  
 7645 GAATAGTGTATGCGGCGACGAGTTGCTCTTGCCCGCGCTCAACACGGGATAATACGCG 7704  
 3308 GAATAGTGTATGCGGCGACGAGTTGCTCTTGCCCGCGCTCAACACGGGATAATACGCG 3367  
 7705 CCACATAGCAGAACTTTAAAGTGTCTCATCTGGAATAAAGCTTCTTCGGGCGGAAACTC 7764  
 3368 CCACATAGCAGAACTTTAAAGTGTCTCATCTGGAATAAAGCTTCTTCGGGCGGAAACTC 3427  
 7765 TCAAGGATCTTACCGCTGTTGAGATCCAGTCCAGTGAAGCCACTCGTCGCCCACTGA 7824  
 3428 TCAAGGATCTTACCGCTGTTGAGATCCAGTCCAGTGAAGCCACTCGTCGCCCACTGA 3487  
 7825 TCTTCAGCATCTTTTACTTTTCCACGAGGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT 7884  
 3488 TCTTCAGCATCTTTTACTTTTCCACGAGGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT 3547  
 7885 GCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTT 7944  
 3548 GCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTT 3607  
 7945 CAATATTTATGAGCATTTATCAGGTTTATGTTCTCATGCGGATACATATTTGAATGT 8004  
 3608 CAATATTTATGAGCATTTATCAGGTTTATGTTCTCATGCGGATACATATTTGAATGT 3667  
 8005 ATTTAGAAAAATAAACAATAGGGGTTCCGGCGCATTTCCCGGAAAGTCCACCTGAC 8064  
 3668 ATTTAGAAAAATAAACAATAGGGGTTCCGGCGCATTTCCCGGAAAGTCCACCTGAC 3727  
 8065 CTCTAGCAAAACCTATTTATCTATCTACCTTAACTTATAAATAAGCGCTATCACGAGGCC 8124

Db 3728 GTCTAAGAAACCATTTATTCATGACATTAACCTTATAAATAAGCGCTATCACGAGGCC 3787  
 QY 8125 TTTCGTCTTCAAGAATTAAATTC 8146  
 Db 3788 TTTCGTCTTCAAGAATTAAATTC 3809

Search completed: September 24, 2004, 10:08:54  
 Job time : 389.227 sess

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 17:04:19 ; Search time 12503.4 seconds  
(without alignments)  
19682.126 Million cell updates/sec

Title: US-10-030-390-3  
Perfect score: 8241  
Sequence: 1 agatctaacatccaagaagc.....aaataacagttattatccg g241

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthm:\*  
3: em\_estim:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055.8	12.8	1070	9 AJ281552	AJ281552 4A3A-P6F1
2	993.2	11.9	1013	12 BM438846	BM438846 IplYr0015
3	924	11.2	1004	9 AJ281480	AJ281480 4A3A-P4G8
4	915.8	11.1	973	14 CD458281	CD458281 Fg08_08h0

C 5	885.2	10.7	917	14	CD458286	CD458286 Fg08_09a0
C 6	872.6	10.6	935	12	BG838279	BG838279 Gc01_10e0
C 7	865	10.5	918	14	CD459092	CD459092 Fg08_08e0
C 8	837.8	10.2	841	9	AL042026	AL042026 DKF2p434E
C 9	829.6	10.1	872	14	CD459085	CD459085 Fg08_08d0
C 10	821.4	10.0	854	12	BM438950	BM438950 IplYr0049
C 11	819.2	9.9	870	14	CD458333	CD458333 Fg08_09e0
C 12	811.2	9.8	1126	28	BZ577702	BZ577702 msh2_5533
C 13	804.8	9.6	819	14	CD649375	CD649375 Gvghd0008
C 14	794.4	9.6	1011	28	BZ576726	BZ576726 msh2_5071
C 15	785.4	9.5	1574	28	BZ576726	BZ576726 msh2_2693
C 16	779.4	9.5	1073	14	CF269652	CF269652 Pcy1C0Ld8
C 17	777.8	9.4	800	9	AJ281449	AJ281449 4A3A-P4D5
C 18	771.6	9.4	1336	28	BZ575810	BZ575810 msh2_4637
C 19	771.2	9.4	1370	28	BZ571721	BZ571721 msh2_2025
C 20	771.2	9.4	789	14	CD280920	CD280920 G44224_42
C 21	770.4	9.3	1249	28	BZ572284	BZ572284 msh2_2572
C 22	763.8	9.3	966	28	BZ570738	BZ570738 msh2_1513
C 23	760	9.2	780	13	BQ825693	BQ825693 1030I29B0
C 24	753.8	9.1	759	14	CD279661	CD279661 G43818_35
C 25	752.2	9.1	1463	28	BZ571475	BZ571475 msh2_1906
C 26	745.4	9.0	844	28	BZ574513	BZ574513 msh2_3706
C 27	745	9.0	786	14	CD458721	CD458721 Fg08_04e0
C 28	741.8	9.0	1003	28	BZ576686	BZ576686 msh2_5053
C 29	740.6	9.0	966	28	BZ575002	BZ575002 msh2_4255
C 30	740.4	9.0	832	12	BQ923768	BQ923768 602825893
C 31	739.2	9.0	914	28	BZ569398	BZ569398 pac82-164
C 32	737.8	9.0	741	14	CD279174	CD279174 G44221_83
C 33	733.2	8.9	811	29	ATH517156	AJ517156 Arabidops
C 34	730.6	8.9	950	28	BZ571129	BZ571129 msh2_1741
C 35	729.2	8.8	998	28	BZ576702	BZ576702 msh2_5060
C 36	728	8.8	954	9	AL044364	AL044364 DKF2p434C
C 37	726.8	8.8	730	14	CD281097	CD281097 G44224_38
C 38	726.2	8.8	1067	9	AU081137	AU081137 AU081137
C 39	726	8.8	840	28	BZ571995	BZ571995 msh2_220.
C 40	725.8	8.8	729	14	CD279546	CD279546 G44222_81
C 41	725.4	8.8	842	12	BI687610	BI687610 603312586
C 42	724.8	8.8	728	14	CD279322	CD279322 G44223_13
C 43	723	8.8	881	28	BZ577222	BZ577222 msh2_5308
C 44	722.8	8.8	726	14	CD281811	CD281811 G44471_30
C 45	722.8	8.8	1394	28	BZ580042	BZ580042 msh2_908.

ALIGNMENTS

RESULT 1	AJ281552	1070 bp	mRNA	linear	EST 30-JUN-2000
AJ281552	4A3A-P6F11-F	Anopheles gambiae	immune competent	4A3A	Anopheles
LOCUS	gambiae cDNA clone 4A3A-P6F11				mRNA sequence.
DEFINITION	AJ281552				
ACCESSION	AJ281552.1	GI:6929432			
VERSION	EST.				
KEYWORDS	Anopheles gambiae (African malaria mosquito)				
SOURCE	Anopheles gambiae				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles				
REFERENCE	1 (bases 1 to 1070)				
AUTHORS	Dionopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Dorochue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B. and Kafatos, F.C.				
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (12)	6619-6624	(2000)	
MEDLINE	20300950				
PUBMED	10841561				
COMMENT	Contact: Dimopoulos G Fotis C. Kafatos laboratory European Molecular Biology Laboratory Meyerhofstrasse 1, 69117 Heidelberg, Germany.				

```

FEATURES
  source
    Location/Qualifiers
      1..1070
        /organism="Anopheles gambiae"
        /mol_type="mRNA"
        /strain="4a r/r"
        /db_xref="taxon:7165"
        /clone="4A3A-P6P11"
        /cell_line="immune competent 4A3A"
        /lab_host="E. coli DH10B"
        /clone_lib="Anopheles gambiae immune competent 4A3A"
        /note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996). Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN
  Query Match
  Best Local Similarity 99.7%; Pred. No. 5e-224; Length 1070;
  Matches 1068; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 6301 TTGCTGGCGTTTTCATAGCTCGGCGCCCTGACGAGCATCACAAAATCGACGCTCA 6360
Db 1 TTGCTGGCGTTTTCATAGCTCGGCGCCCTGACGAGCATCACAAAATCGACGCTCA 60

QY 6361 AGTCAGAGGTGGGAAACCCGACAGGACTATAAGATACACAGGCGTTTCCCGCTGGAAGC 6420
Db 61 AGTCAGAGGTGGGAAACCCGACAGGACTATAAGATACACAGGCGTTTCCCGCTGGAAGC 120

QY 6421 TCCCTCGTGCCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGGCTTTCTTC 6480
Db 121 TCCCTCGTGCCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGGCTTTCTTC 180

QY 6481 CTTTCGGGAAGCGGGCGTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAG 6540
Db 181 CTTTCGGGAAGCGGGCGTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAG 240

QY 6541 GTGCTTGCTCAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGCCGACCGCTGCGCC 6600
Db 241 GTGCTTGCTCAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGCCGACCGCTGCGCC 300

QY 6601 TTATCCGGTAACCTATCGTCTTGTAGTCCAAACCCGGTAAGACACACACTTATCGCACTGGCA 6660
Db 301 TTATCCGGTAACCTATCGTCTTGTAGTCCAAACCCGGTAAGACACACACTTATCGCACTGGCA 360

QY 6661 CGAGCCACTGTACAGATTAGCAGAGCAGGTATGTAGGCGGTGCTACAGATTCTTG 6720
Db 361 CGAGCCACTGTACAGATTAGCAGAGCAGGTATGTAGGCGGTGCTACAGATTCTTG 420

QY 6721 AAGTGGTGGCTAACTACGCTACACTAGAGACAGTATTTGGTATCTGCGCTCTGCTG 6780
Db 421 AAGTGGTGGCTAACTACGCTACACTAGAGACAGTATTTGGTATCTGCGCTCTGCTG 480

QY 6781 AAGCCAGTTACCTCGGAAAGAGTTGGTGTAGTCTTGTATCGGCGAAACAAACCCCGCT 6840
Db 481 AAGCCAGTTACCTCGGAAAGAGTTGGTGTAGTCTTGTATCGGCGAAACAAACCCCGCT 540

QY 6841 GGTAGCGGTGGTTTTTTTGTGTAAGCAGCATACGCGCAGAAAAAAGGATCTCAA 6900
Db 541 GGTAGCGGTGGTTTTTTTGTGTAAGCAGCATACGCGCAGAAAAAAGGATCTCAA 600

QY 6901 GAAATGCTTTGATCTTTTCTACGGGTCTGAGCTCAGTGGACGAAACATCAGTTAA 6960
Db 601 GAAATGCTTTGATCTTTTCTACGGGTCTGAGCTCAGTGGACGAAACATCAGTTAA 660

QY 6961 GGGATTTTGGTCATGAGATTCAAAAAGGATCTTCACCTAGATCTCTTTAAATTAATAA 7020
Db 661 GGGATTTTGGTCATGAGATTCAAAAAGGATCTTCACCTAGATCTCTTTAAATTAATAA 720

QY 7021 TGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTTACCAATGC 7080

```

```

Db 721 TGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTTACCAATGC 780
QY 7081 TTAATCAGTGAAGCACCCTATCTCAGCGATCTGTCTATTTTCGTTTCATCATAGTTCCCTGA 7140
Db 781 TTAATCAGTGAAGCACCCTATCTCAGCGATCTGTCTATTTTCGTTTCATCATAGTTCCCTGA 840
QY 7141 CTCCCGCTCTGTAGTAACTACGATACGGAGGGCTTACCATCTGGCCCGAGTCTGCA 7200
Db 841 CTCCCGCTCTGTAGTAACTACGATACGGAGGGCTTACCATCTGGCCCGAGTCTGCA 900
QY 7201 ATGATACCGGAGACCCACCGCTCACCGGCTCCAGATTATATCAGCAATAAACCCAGCCGCC 7260
Db 901 ATGATACCGGAGACCCACCGCTCACCGGCTCCAGATTATATCAGCAATAAACCCAGCCGCC 960
QY 7261 GGAAGGCCCGGACCGGAGAGTGTCTGCAATTTATCGGCTCCATCCAGTCTATTAAT 7320
Db 961 GGAAGGCCCGGACCGGAGAGTGTCTGCAATTTATCGGCTCCATCCAGTCTATTAAT 1019
QY 7321 TGTTCGGGGAAGCTAGAGTAACTAGTTTCGCCAGTTAATAGTTTGGCAAC 7371
Db 1020 TGTTCGGGGAAGCTAGAGTAACTAGTTTCGCCAGTTAATAGTTTGGCAAC 1070

RESULT 2
BM438846 1013 bp mRNA linear EST 31-JAN-2002
DEFINITION IPIV00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
sequence.
ACCESSION BM438846
VERSION BM438846.1 GI:18460568
KEYWORDS EST.
SOURCE Ictalurus punctatus (channel catfish)
ORGANISM Ictalurus punctatus
REFERENCE 1 (bases 1 to 1013)
AUTHORS Feng J., Kucuktas H., Kocabas A., Li P. and Liu Z.
TITLE Transcriptome of channel catfish (Ictalurus punctatus): initial
analysis of expressed sequence tags from the liver
JOURNAL Unpublished (2002)
COMMENT Contact: Liu Z
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University, Auburn, AL 36849, USA
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
Seq primer: M13 Reverse.
FEATURES
  source
    Location/Qualifiers
      1..1013
        /organism="Ictalurus punctatus"
        /mol_type="mRNA"
        /db_xref="taxon:7998"
        /clone_lib="Liver cDNA library"
        /note="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: SalI"

ORIGIN
  Query Match
  Best Local Similarity 99.3%; Pred. No. 7.5e-208; Length 1013;
  Matches 1007; Conservative 1; Mismatches 4; Indels 2; Gaps 2;

QY 6874 ATTACGCGCAGAAAAAGGATCTCAAGAAGATCTTTTGATCTTTTCTACGGGTCTGAC 6933
Db 1013 ATTACGCGCAGAAAAAGGATCTCAAGAAGATCTTTTGATCTTTTCTACGGGTCTGAC 954
QY 6934 GCTCAGTGAACGAAACCTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAGGATC 6993
Db 953 GCTCAGTGAACGAAACCTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAGGATC 894

```





Db 66 ATAGGGGTTCCGGCACAATTTCCTCCGAAASTGCCACTAAATGTGAAG 18

RESULT 5  
CD458286/c  
LOCUS  
DEFINITION Fg08\_09a03\_R Fg08\_AAFC\_ECORC\_Fusarium\_graminearum\_complex\_substrate  
EST 03-JUN-2003  
CD458286  
Gibberella zeae cDNA clone Fg08\_09a03, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

CD458286 917 bp -mRNA linear  
Fg08\_09a03\_R Fg08\_AAFC\_ECORC\_Fusarium\_graminearum\_complex\_substrate  
Gibberella zeae cDNA clone Fg08\_09a03, mRNA sequence.  
CD458286 GI:31373026  
EST.  
Gibberella zeae  
Gibberella zeae  
Eukaryota: Fungi; Ascomycota: Pezizomycotina; Sordariomycetes;  
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
1 (bases 1 to 917)  
Watson,R.J., Heys,R., Chapados,J., Couroux,P., Harris,L.J.,  
Hattori,J.J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A.,  
Spott,D. and Tinker,N.A.  
A cDNA library prepared from Fusarium graminearum grown on a  
complex plant substrate  
Unpublished (2003)  
Contact: Watson, Robert.J.  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-Food Canada  
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
CANADA  
Tel: (613) 759-1655  
Fax: (613) 759-1701  
Email: watsonrj@agr.gc.ca.

FEATURES  
source  
1. 917  
/organism="Gibberella zeae"  
/mol\_type="mRNA"  
/strain="DAOM 180378"  
/db\_xref="taxon:5518"  
/clone="Fg08\_09a03"  
/tissue\_type="Mycelium"  
/dev\_stage="Asexual"  
/lab\_host="E. coli DH10B"  
/clone\_lib="Fg08\_AAFC\_ECORC\_Fusarium\_graminearum\_complex\_1  
ubstrate"  
/notes="Vector: pBluescript II+; Site 1: EcoRI; Site 2:  
XhoI; Fusarium graminearum grown on a complex plant  
substrate-- wheat leaves treated to remove most of the low  
molecular weight, water-soluble components."

ORIGIN  
Query Match 10.7%; Score 885.2; DB 14; Length 917;  
Best Local Similarity 99.1%; Pred. No. 5.1e-186;  
Matches 890; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 7174 GCGTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACGGGTCCA 7233  
Db 917 GGGTTACCACTTGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACGGGTCCA 858  
QY 7234 GATTTATCAGCAATAACACGACGAGCGGAGCGGAGCGGAGAGTGGTCTCGCAACT 7293  
Db 857 GATTTATCAGCAATAACCCAGCCAGCGGAGGCGGAGGCGAGAGTGGTCTCGCAACT 798  
QY 7294 TTATCGCGCTCCATCCAGTCTATTAAATTGTTGCGGGAAGCTAGAGTAGTTCGCCA 7353  
Db 797 TTATCGCGCTCCATCCAGTCTATTAAATTGTTGCGGGAAGCTAGAGTAGTTCGCCA 738  
QY 7354 GTTAAATAGTTTGGCGCAACGTTGTTGCAATGCTGCAGGCACTCTGTGTGCACGCTCGTCG 7413  
Db 737 GTTAAATAGTTTGGCAACGTTGTTGCCATGCTCAGGCACTCTGTGTGCACGCTCGTCG 678  
QY 7414 TTTGGTAGTGCCTCATTTACGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC 7473  
Db 677 TTTGGTAGTGCCTCATTTACGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC 618



```

/mol_type="mRNA"
/cultivar="1035"
/db_xref="taxon:45687"
/clone="Gc01_10e07"
/tissue_type="Leaves, stem"
/clone_lib="Gc01_AAPC_ECORC_cold_stressed_glycine_clandest
ina"
/notes="Vector: Bluescript SK-/XhoI-EcoRI; Site 1: EcoRI;
Site 2: XhoI; Plants incubated at 2 degrees under 12 hours
of light/day. Harvested after only 2-3 days of cold
treatment. cDNA was prepared with the Uni-Zap cDNA kit
from Stratagene. Eco RI adapters were linked followed by
digest with Xho I/Eco RI and ligated to pBluescript."

ORIGIN
Query Match          10.6%; Score 872.6; DB 12; Length 935;
Best Local Similarity 97.5%; Pred. No. 3.3e-183;
Matches 880; Conservative 14; Mismatches 8; Indels 1; Gaps 1;

QY 7170 GGAGGGCTTACATCTGGCCCGCAGTGTGCAATGATACCGCGAGACCCAGCTCACCGGC 7229
DB 928 GGGGGCTTACCATCTGCCCCAGKGTGCAATGATACCGGAGGCCAGCTCMCCGC 869
QY 7230 TCCAGATTATCAGCAATAAACAGCAGCAGCGGAGGGCGGAGAGTGGTCTCTGC 7289
DB 868 YCCAGATTATCAGCAATAAACAGCAGCAGCGGAGGGCGGAGAGTGGTCTCTGC 809
QY 7290 AACTTTATCCGCTCCATCTCTATTATTTGTCGGGGAAGCTAGAGTAGTAGTTC 7349
DB 808 AKCTTTATCCGCTCCATCTCTATTATTTGTCGGGGAAGCTAGAGTAGTAGTTC 749
QY 7350 GCCAGTTAATAGTTTGGCAACGTTTGGCAATGCTGAGGAGTGTGGTGTGTCAGCTC 7409
DB 748 GCCAGTTAATAGTTTGGCAACGTTTGGCAATGCTGAGGAGTGTGGTGTGTCAGCTC 689
QY 7410 GTCTGTTTGGTATGGCTTCATTCAGCTCGGTTCCCAACGATCAAGCGAGTTACATGATC 7469
DB 688 GTCTGTTTGGTATGGCTTCATTCAGCTCGGTTCCCAACGATCAAGCGAGTTACATGATC 629
QY 7470 CCCCATGTTGCAAAAAGGGTTAGCTCTCTGGTCTCCGATCGTTGTGCAAGATGA 7529
DB 628 CCCCATGTTGCAAAAAGGGTTAGCTCTCTGGTCTCCGATCGTTGTGCAAGATGA 569
QY 7530 GTT-GGCCGCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCA 7588
DB 568 GTTGGCGCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCA 509
QY 7589 TGCCATCGTAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATTCAGAT 7648
DB 508 TGCCATCGTAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATTCAGAT 449
QY 7649 AGTGATCGCGCGACCGAGTTGCTCTTGGCCGCGGTCAACACGGGATAATACGGCGCAC 7708
DB 448 AGTGATCGCGCGACCGAGTTGCTCTTGGCCGCGGTCAATACGGGATAATACGGCGCAC 389
QY 7709 ATAGCAGAACTTAAAGTGTCTCATCTTGGAAAGCTTCTCGGGCGAAACTCTCAA 7768
DB 388 ATAGCAGAACTTAAAGTGTCTCATCTTGGAAAGCTTCTCGGGCGAAACTCTCAA 329
QY 7769 GGATCTTACCCTGTTGAGATTCAGTTTCGATGTAAACCCACTGTGCAACCAACTGATCTT 7828
DB 328 GGATCTTACCCTGTTGAGATTCAGTTTCGATGTAAACCCACTGTGCAACCAACTGATCTT 269
QY 7829 CAGCATCTTTTACCTTACAGCGTTTCTGCGTGAGCAAAACAGGAAGCAAAATGCG 7888
DB 268 CAGCATCTTTTACCTTACAGCGTTTCTGCGTGAGCAAAACAGGAAGCAAAATGCG 209
QY 7889 CAAAAAGGGGAATAAGGGCGACACGGAAATGTTGAATCTCATCTCTCTTTTTCAT 7948
DB 208 CAAAAAGGGGAATAAGGGCGACACGGAAATGTTGAATCTCATCTCTCTTTTTCAT 149
QY 7949 ATTATGAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGAATGATTT 8008

```

```

DB 148 ATTATTGAACATTATATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGATTT 89
QY 8009 AGAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTCT 8068
DB 88 AGAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGAAAGTGCACCTGATTTGT 29
QY 8069 AAG 8071
DB 28 AAG 26

RESULT 7
CD459092/c
LOCUS          918 bp      mRNA      linear      EST 03-JUN-2003
DEFINITION    Fg08_08e02_R Fg08 AAPC ECORC Fusarium graminearum_complex_substrate
Gibberella zeae cDNA clone Fg08_08e02, mRNA sequence.
ACCESSION     CD459092
VERSION       CD459092.1 GI:31373832
KEYWORDS      EST.
SOURCE        Gibberella zeae
ORGANISM      Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE     1 (bases 1 to 918)
AUTHORS       Watson, R.J., Heys, R., Chapados, J., Couroux, P., Harris, L.J.,
Hattori, J., Lacroix, C., Ouellet, T., Robert, L.S., Singh, J.A.,
Sprott, D. and Tinker, N.A.
TITLE         A cDNA library prepared from Fusarium graminearum grown on a
              complex plant substrate
              Unpublished (2003)
JOURNAL       Contact: Watson, Robert.J.
COMMENT       Eastern Cereal and Oilseed Research Centre
              Agriculture and Agri-Food Canada
              Bldg. 20, Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
              CANADA
              Tel: (613) 759-1655
              Fax: (613) 759-1701
              Email: watsonrj@agr.gc.ca.

FEATURES             source
                     1..918
                        /organism="Gibberella zeae"
                        /mol_type="mRNA"
                        /strain="DAOM 180378"
                        /db_xref="taxon:5518"
                        /clone="Fg08_08e02"
                        /tissue_type="Mycelium"
                        /dev_stage="Asexual"
                        /lab_host="F. coli DH10B"
                        /clone_lib="Fg08_AAPC_ECORC_Fusarium_graminearum_complex_s
                        ubstrate"
                        /note="Vector: pBluescript II+; Site 1: EcoRI; Site 2:
                        XhoI; Fusarium graminearum grown on a complex plant
                        substrate-- wheat leaves treated to remove most of the low
                        molecular weight, water-soluble components."

ORIGIN
Query Match          10.5%; Score 865; DB 14; Length 918;
Best Local Similarity 98.8%; Pred. No. 1.6e-181;
Matches 890; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

QY 7173 GGGCTTACCATCTGGCCCCAGTGTGCAATGA-TACCCGAGAGACCCAGCTCACCGGCTC 7231
DB 918 GGGCTTACCATCTGGCCCCAGKGTGCAATGATTTACCGGAGMCCCGAGCTCACCGGCTC 859
QY 7232 CAGATTATCAGCAATAAACAGCAGCAGCGGAGGGCGGAGAGTGGTCTCTCAA 7291
DB 858 CAGATTATCAGCAATAAACAGCAGCAGCGGAGGGCGGAGAGTGGTCTCTCAA 799
QY 7292 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCTAGAGTAGTAGTTCGC 7351
DB 798 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCTAGAGTAGTAGTTCGC 739
QY 7352 CAGTTAATAGTTTGGCAACGTTGTTGCCATTGTCGAGGCAATCGTGGTGTCAAGCTCGT 7411

```

Db 738 CAGTTAATAGTTGGCGCAAGCTTGTGGCAATGCTACAGGCATCGTGGTGTACGCTCGT 679  
Qy 7412 CGTTTGTATGGCTTCAATTCAGCTCCGGTCCCAACGATCAAGCGAGTTACATGATCCC 7471  
Db 678 CGTTTGTATGGCTTCAATTCAGCTCCGGTCCCAACGATCAAGCGAGTTACATGATCCC 619  
Qy 7472 CCATGTTGTGCAAAAAGCGGTAGCTCCCTTCGGTCTCCGATCGTTGTGTCAGAAAGTAAGT 7531  
Db 618 CCATGTTGTGCAAAAAGCGGTAGCTCCCTTCGGTCTCCGATCGTTGTGTCAGAAAGTAAGT 559  
Qy 7532 TGGCCGAGTGTATCACTCA-TGGTTATGGCAGCACTGCATATATCTCTTACTGTCAATG 7590  
Db 558 TGGCCGAGTGTATCACTCACTGTTGATGGCAGCACTGCATATATCTCTTACTGTCAATG 499  
Qy 7591 CCATCCGTAAAGATCTTTCTGTGACTGGTGAGTACTCAACCAAGTCATCTCAGATAG 7650  
Db 498 CCATCCGTAAAGATCTTTCTGTGACTGGTGAGTACTCAACCAAGTCATCTCAGATAG 439  
Qy 7651 TGTATGCGGCGACCGAGTCTCTCTTGGCCGGCGTCAACAGCGGATATACCGGCGCACAT 7710  
Db 438 TGTATGCGGCGACCGAGTCTCTCTTGGCCGGCGTCAATACGGGATATACCGGCGCACAT 379  
Qy 7711 AGCAGAACTTTAAAGTGTCTCATCTATGGAAGCTTCTTCGGGGGGAAGTCTCAAGG 7770  
Db 378 AGCAGAACTTTAAAGTGTCTCATCTATGGAAGCTTCTTCGGGGGGAAGTCTCAAGG 319  
Qy 7771 ATCTTACCGCTGTGTAGATCCAGTTCATGTCAACCCACTCGTGCACCCAACTGATCTTCA 7830  
Db 318 ATCTTACCGCTGTGTAGATCCAGTTCATGTCAACCCACTCGTGCACCCAACTGATCTTCA 259  
Qy 7831 GCATCTTTTACTTTCACAGCGTTCCTGGGTGAGCAAAACAGAGAAAGGCAAAATGCCGA 7890  
Db 258 GCATCTTTTACTTTCACAGCGTTCCTGGGTGAGCAAAACAGAGAAAGGCAAAATGCCGA 199  
Qy 7891 AAAAAGGGAATAAGGGGACACGGAATGTTGAATCTACTACTCTCTCTTTTCAATAT 7950  
Db 198 AAAAAGGGAATAAGGGGACACGGAATGTTGAATCTACTACTCTCTCTTTTCAATAT 139  
Qy 7951 TATTGAAGCAATTATCAGGTTATTTCTCATAGCGGATACATATTTGAATGTAATTAG 8010  
Db 138 TATTGAAGCAATTATCAGGTTATTTCTCATAGCGGATACATATTTGAATGTAATTAG 79  
Qy 8011 AAAAATAACAAATAGGGTTCGCGCACATTTTCCGGAAGTGCACCTGACGCTTAA 8070  
Db 78 AAAAATAACAAATAGGGTTCGCGCACATTTTCCGGAAGTGCACCTGACCTAAATGTAA 19  
Qy 8071 G 8071  
Db 18 G 18

RESULT 8  
AL042026/c  
LOCUS  
DEFINITION  
DEFINITION DKFZp434E111 r1 434 (synonym: htes3) Homo sapiens cDNA clone  
ACCESSION  
VERSION AL042026.1 GI:5421372  
KEYWORDS  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 841)  
AUTHORS Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
TITLE EST (Poustka, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: MIPS  
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ): Email s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No si sequence available.  
This clone (DKFZp434E111) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source  
1..841  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp434E111"  
/tissue\_type="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="434 (synonym: htes3)"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"

ORIGIN  
Query Match 10.2%; Score 837.8; DB 9; Length 841;  
Best Local Similarity 99.8%; Pred No. 1.8e-175;  
Matches 339; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 7185 TGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCGCTCCAGATTTATCAGC 7244  
Db 841 TGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCGCTCCAGATTTATCAGC 782  
Qy 7245 AATAACACAGCCAGCGGAGGCGGAGAGTGGTCTGCAACTTTATCGGCTC 7304  
Db 781 AATAACACAGCGGAGGCGGAGAGTGGTCTGCAACTTTATCGGCTC 722  
Qy 7305 CATCCAGTCTATTAAATTGTTCCGCGGAAGCTAGAGTAGTTCGCCAGTTAATAGTTT 7364  
Db 721 CATCCAGTCTATTAAATTGTTCCGCGGAAGCTAGAGTAGTTCGCCAGTTAATAGTTT 662  
Qy 7365 GCGCAACGTTGTGGCAATGCTGAGGCATCGTGGTGTACGCTCGTGGTGGTATGCG 7424  
Db 661 GCGCAACGTTGTGGCAATGCTGAGGCATCGTGGTGTACGCTCGTGGTGGTATGCG 602  
Qy 7425 TTCAATTCAGTCCGGTTCCTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCA 7484  
Db 601 TTCAATTCAGTCCGGTTCCTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCA 542  
Qy 7485 AAAAGCGTTAGTCTCTTCGCTTCGATCGTTGTGCAAGTAGTAAAGTGGCCGAGTGT 7544  
Db 541 AAAAGCGTTAGTCTCTTCGCTTCGATCGTTGTGCAAGTAGTAAAGTGGCCGAGTGT 482  
Qy 7545 ATCACTCATGTTATGGCAGCACTGCATTAATTCCTTACTCTCATGCTCATCGCTAAGATG 7604  
Db 481 ATCACTCATGTTATGGCAGCACTGCATTAATTCCTTACTCTCATGCTCATCGCTAAGATG 422  
Qy 7605 CTTTCTCTGACTGGTGAAGTACTCAACCAAGTCAATTCCTGAGAATAGTGTATCGCGGAC 7664  
Db 421 CTTTCTCTGACTGGTGAAGTACTCAACCAAGTCAATTCCTGAGAATAGTGTATCGCGGAC 362  
Qy 7665 GAGTTGCTTCTTCGCGGCTCAACACGGGATTAATACCGGCGCATACAGCAACTTTAAA 7724  
Db 361 GAGTTGCTTCTTCGCGGCTCAATACGGGATTAATACCGGCGCATACAGCAACTTTAAA 302  
Qy 7725 AGTGTCTCATCTATGGAACCGTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGT 7784  
Db 301 AGTGTCTCATCTATGGAACCGTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGT 242  
Qy 7785 GAGATCCAGTTGAGTAAACCACTCGTGGCAACCAAGTATCTTCAGCATCTTTTACTTT 7844  
Db 241 GAGATCCAGTTGAGTAAACCACTCGTGGCAACCAAGTATCTTCAGCATCTTTTACTTT 182  
Qy 7845 CACCAGCGTTTCTCGGTGAGCAAAACAGGAGGCAAAATCCCGCAAAAGGGAATAAG 7904  
Db 181 CACCAGCGTTTCTCGGTGAGCAAAACAGGAGGCAAAATCCCGCAAAAGGGAATAAG 122

```

QY 7905 GCGACACGGAATGTTGAATACATACATCTTCTTCTTCAATATATTATTAAGACATTTA 7964
|
Db 121 GCGACACGGAATGTTGAATACATACATCTTCTTCTTCAATATATTATTAAGACATTTA 62
|
QY 7965 TCAGGTTATTGTTCTCATAGCGGATACATATTTGAATATTGAAATTAAGAAAATAACAAAT 8024
|
Db 61 TCAGGTTATTGTTCTCATAGCGGATACATATTTGAATATTGAAATTAAGAAAATAACAAAT 2
|
QY 8025 A 8025
|
Db 1 A 1

RESULT 9
CD459085/c
LOCUS
DEFINITION Fg08_08d03 R Fg08_AAFc ECORC Fusarium graminearum complex_substrate
Gibberella_zeae cDNA clone Fg08_08d03, mRNA sequence.
ACCESSION CD459085
VERSION CD459085.1 GI:31373825
KEYWORDS EST.
SOURCE Gibberella zeae
ORGANISM Gibberella zeae
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
AUTHORS Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J.,
Hattori,J., Lacroix,C., Ouellet,F., Robert,L.S., Singh,J.A.,
Spott,D. and Tinker,N.A.
TITLE A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
JOURNAL Unpublished (2003)
COMMENT Contact: Watson, Robert.J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Blg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1655
Fax: (613) 759-1701
Email: watsonrj@agr.gc.ca.

FEATURES
Location/Qualifiers
1..872
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="Fg08_08d03"
/tissue_type="Mycelium"
/dev stage="Asexual"
/lab_hosts="E. coli DH10B"
/clone_lib="Fg08_AAFc_ECORC_Fusarium_graminearum_complex_s
ubstrate"
/notes="vector: pBluescript II+; Site 1: EcoRI; Site 2:
XhoI; Fusarium graminearum grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."

ORIGIN
Query Match 10.18; Score 829.6; DB 14; Length 872;
Best Local Similarity 98.9; Pred. No. 1.2e-173;
Matches 829; Conservative 5; Mismatches 43; Indels 0; Gaps 0;

QY 7215 CCCAGCTCACCGGCTCCAGATTATACGCAATAAACACCGAGCGGAGCGGAGCG 7274
|
Db 872 CCCAGCTCACCGGCTCCAGATTATACGCAATAAACACCGAGCGGAGCGGAGCG 813
|
QY 7275 CAGAGTGGTCCCTGACATTTATCCGCCCTCATCCAGTCTATTAATTTGCGCGGAGC 7334
|
Db 812 CAGAGTGGTCCCTGACATTTATCCGCCCTCATCCAGTCTATTAATTTGTTGCGGAGC 753
|
QY 7335 TAGAGTAAGTAGTTGCGCAGTTAATAGTTTGCACACGTTTGTGTCATTTGTCAGGAT 7394
|
Db 752 TAGAGTAAGTAGTTGCGCAGTTAATAGTTTGCACACGTTTGTGTCATTTGTCAGGAT 693

```

```

QY 7395 CGTGGTGTCAAGCTCGTCTGTTGGTATGGCTTCATTTCAGCTCCGTTCCCAACGATCAAG 7454
|
Db 692 CGTGGTGTCAAGCTCGTCTGTTGGTATGGCTTCATTTCAGCTCCGTTCCCAACGATCAAG 633
|
QY 7455 GCGAGTTACATGATCCCCCATGTTGTCAAAAAGCGTTAGTCTCTTCGTTCCGAT 7514
|
Db 632 GCGAGTTACATGATCCCCCATGTTGTCAAAAAGCGTTAGTCTCTTCGTTCCGAT 573
|
QY 7515 CGTTGTCAAGTAAGTTGCGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAA 7574
|
Db 572 CGTTGTCAAGTAAGTTGCGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAA 513
|
QY 7575 TTCTTCTACTGTTCATCCCATCCGTAAAGATGCTTTTCTGTGACTGGTACTCAACCAA 7634
|
Db 512 TTCTTCTACTGTTCATCCCATCCGTAAAGATGCTTTTCTGTGACTGGTACTCAACCAA 453
|
QY 7635 GTCATTTCTGAGATAGTGTATGCGCGCAGCGAGTTGCTCTTCCCGCGGTCAACACGGA 7694
|
Db 452 GTCATTTCTGAGATAGTGTATGCGCGCAGCGAGTTGCTCTTCCCGCGGTCAACACGGA 393
|
QY 7695 TAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAACCTTTCTTCGG 7754
|
Db 392 TAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAACCTTTCTTCGG 333
|
QY 7755 GCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCCGATGTAAACCACTCGTGC 7814
|
Db 332 GCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCCGATGTAAACCACTCGTGC 273
|
QY 7815 ACCCACTGATCTTCAGCATCTTTTACTTTTACCAGCTTTTCTGGTGAGCAAAAACAGG 7874
|
Db 272 ACCCACTGATCTTCAGCATCTTTTACTTTTACCAGCTTTTCTGGTGAGCAAAAACAGG 213
|
QY 7875 AAGGCAAAATGCGCAAAAAGGAATAGGCGCACGCGAAATGTTGAATACTCACTACT 7934
|
Db 212 AAGGCAAAATGCGCAAAAAGGAATAGGCGCACGCGAAATGTTGAATACTCACTACT 153
|
QY 7935 CTTTCTTTTCAATATATTATTAAGCATTTATCAGGTTTATTGTCTCATGAGCGGATACAT 7994
|
Db 152 CTTTCTTTTCAATATATTATTAAGCATTTATCAGGTTTATTGTCTCATGAGCGGATACAT 93
|
QY 7995 ATTTGATGTATTAGAAAATAAACAAATAGGGTTCCGCGCATTTTCCCGGAAA 8052
|
Db 92 ATTTGATGTATTAGAAAATAAACAAATAGGGTTCCGCGCATTTTCCCGGAAA 35

RESULT 10
BM438950/c
LOCUS
DEFINITION BM438950 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
Sequence.
ACCESSION BM438950
VERSION BM438950.1 GI:18460672
KEYWORDS EST.
SOURCE Ictalurus punctatus (channel catfish)
ORGANISM Ictalurus punctatus
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
1 (bases 1 to 854)
Feng,J., Kucuktas,H., Kocabas,A., Li,P. and Liu,Z.
Transcriptome of channel catfish (Ictalurus punctatus): initial
analysis of expressed sequence tags from the liver
Unpublished (2002)
Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu

```

Db	15	AGGAGGCAAAATGC 1
Seq primer: M13 Reverse.		
Location/Qualifiers		
1. .854		
/organism="Ictalurus punctatus"		
/mol_type="mRNA"		
/db_xref="taxon:7998"		
/clone_lib="liver cDNA library"		
/note="Organ: Liver; Vector: pSport1; Site_1: NotI; Site_2: SalI"		
ORIGIN		
QY	7033	TCAATCTAAAGTATATAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 7092
Db	854	TCATCTCAGCATATATAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 795
QY	7093	GCACCTATCTCAGGATCTGTCTATTTTCGTTTCATCCATAGTTCGCTGACTCCCGTCGTG 7152
Db	794	GCACCTATCTCAGGATCTGTCTATCTCGTTTCATCCATAGTTCGCTGACTCCCGTCGTG 735
QY	7153	TAGATACTACGATACGGAGGCTTACCATCTGGCCCACTGCTGCAATGATACCGCA 7212
Db	734	TAGATACTACGATACGGAGGCTTACCATCTGGCCCACTGCTGCAATGATACCGCA 675
QY	7213	GACCCAGCTCAGCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCGAG 7272
Db	674	GACCCAGCTCAGCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCGAG 615
QY	7273	CGCAGAGTGTCTCTGCACTTTATCCGCTCCATCCAGTCTATTAATGTTCCCGGAA 7332
Db	614	CGCAGAGTGTCTCTGCACTTTATCCGCTCCATCCAGTCTATTAATGTTCCCGGAA 555
QY	7333	GCTAGAGTAAGTATGTCGCCAGTTAATAGTTTTCGCCAAACGTTTTCGCCATTTGTCGAGGC 7392
Db	554	GCTAGAGTAAGTATGTCGCCAGTTAATAGTTTTCGCCAAACGTTTTCGCCATTTGTCGAGGC 495
QY	7393	ATGCTGTGTTCACGCTCGTCTGTTGTTGATGCTTCAATCAGTCCGGTTCGCCAAGCATCA 7452
Db	494	ATGCTGTGTTCACGCTCGTCTGTTGTTGATGCTTCAATCAGTCCGGTTCGCCAAGCATCA 435
QY	7453	AGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCCTTCGTCCTCCG 7512
Db	434	AGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCCTTCGTCCTCCG 376
QY	7513	ATCGTTGTCTAGAGTAAGTTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGCAT 7572
Db	375	ATCGTTGTCTAGAGTAAGTTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGCAT 316
QY	7573	AATTCCTTACTGTCACTCCGTAAGTGTCTTTCTGTGACTGTTGAGTACTCAACC 7632
Db	315	AATTCCTTACTGTCACTCCGTAAGTGTCTTTCTGTGACTGTTGAGTACTCAACC 256
QY	7633	AAGTCATTCTGAGATAGTGTATCGCGGACCGAGTTGCTTTGCCCGGCGTCAACACGG 7692
Db	255	AAGTCATTCTGAGATAGTGTATCGCGGACCGAGTTGCTTTGCCCGGCGTCAACACGG 196
QY	7693	GATAATACCGCGGCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAAAAGCTTCTTCG 7752
Db	195	GATAATACCGCGGCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAAAAGCTTCTTCG 136
QY	7753	GGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA-CCCACTCG 7811
Db	135	GGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAAGCCACTCG 76
QY	7812	TGCACCAACTGATCTTCAGCATCTTTTACTTTTACACAGGTTTCTGGGTGAGCAAAAAC 7871
Db	75	TGCACCAACTGATCTTCAGCATCTTTTACTTTTACACAGGTTTCTGGGTGAGCAAAAAC 16
QY	7872	AGGAAGGCAAAATGC 7885
Query Match 9.9%; Score 819.2; DB 14; Length 870;		
Best Local Similarity 98.7%; Pred. No. 2.6e-171;		
Matches 843; Conservative 3; Mismatches 6; Indels. 2; Gaps 2;		
QY	7219	CGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCGAGCGAGA 7278
Db	870	CGCTCVCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCGAGCGAGA 811
QY	7279	AGTGTCCTGCACTTTATCCGCTCCATCGATCTATTATTGTTGCCGGGAAGCTAGA 7338
Db	810	AGTGTCCTGCACTTTATCCGCTCCATCCAGTCTATTAAATGTTGCCGGGAAGCTAGA 751
QY	7339	GTAAGTAGTTCGCAAGTTAATAGTTTGCACAAAGTTGTTCATTTGTCAGGCAATCGTG 7398
Db	750	GTAAGTAGTTCGCAAGTTAATAGTTTGCACAAAGTTGTTCATTTGTCAGGCAATCGTG 691
QY	7399	GTGTACAGTCTGCTGTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGCA 7458
Db	690	GTGTACAGTCTGCTGTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGCA 631
QY	7459	GTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCCTTCGTTCCCGCATCGTT 7518
Db	630	GTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCCTTCGTTCCCGCATCGTT 572
QY	7519	GTGAG-AGGTAAGTTGGCGGAGTGTATGATCTCATGTTATGGCAGCACTGCATAATTC 7577
/organism="Gibberella zeae"		
/mol_type="mRNA"		
/strain="DAOM 180378"		
/db_xref="taxon:5518"		
/clone_lib="Fg08_09e07"		
/tissue_type="Mycelium"		
/dev_stage="Asexual"		
/lab_host="E. coli DH10B"		
/clone_lib="Fg08_AAFc_ECORC_Fusarium_graminearum_complex_s- substrate"		
/note="Vector: pBluescript II+; Site_1: EcoRI; Site_2: XhoI; Fusarium graminearum grown on a complex plant substrate-- wheat leaves treated to remove most of the low molecular weight, water-soluble components."		
ORIGIN		
QY	7219	CGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCGAGCGAGA 7278
Db	870	CGCTCVCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCGAGCGAGA 811
QY	7279	AGTGTCCTGCACTTTATCCGCTCCATCGATCTATTATTGTTGCCGGGAAGCTAGA 7338
Db	810	AGTGTCCTGCACTTTATCCGCTCCATCCAGTCTATTAAATGTTGCCGGGAAGCTAGA 751
QY	7339	GTAAGTAGTTCGCAAGTTAATAGTTTGCACAAAGTTGTTCATTTGTCAGGCAATCGTG 7398
Db	750	GTAAGTAGTTCGCAAGTTAATAGTTTGCACAAAGTTGTTCATTTGTCAGGCAATCGTG 691
QY	7399	GTGTACAGTCTGCTGTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGCA 7458
Db	690	GTGTACAGTCTGCTGTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGCA 631
QY	7459	GTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCCTTCGTTCCCGCATCGTT 7518
Db	630	GTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCCTTCGTTCCCGCATCGTT 572
QY	7519	GTGAG-AGGTAAGTTGGCGGAGTGTATGATCTCATGTTATGGCAGCACTGCATAATTC 7577

Db	15	AGGAAGGCAAAATGC 1
RESULT 11		
CD458333/c		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
1. .870		
/organism="Gibberella zeae"		
/mol_type="mRNA"		
/strain="DAOM 180378"		
/db_xref="taxon:5518"		
/clone="Fg08_09e07"		
/tissue_type="mycelium"		
/dev_stage="Asexual"		
/lab_host="E. coli DH10B"		
/clone_lib="Fg08_AAFCEORC_Fusarium_graminearum_complex_s"		
ubstrate"		
/note=Vector: pBluescript II+; Site 1: EcoRI; Site 2:		
XhoI; Fusarium graminearum grown on a complex plant		
substrate-- wheat leaves treated to remove most of the low		
molecular weight, water-soluble components."		
ORIGIN		
Query Match		
Best Local Similarity 9.9%; Score 819.2; DB 14; Length 870;		
Matches 843; Conservative 3; Mismatches 6; Indels 2; Gaps 2;		
QY	7219	CGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGAGCGGCGAGCAGA 7278
Db	870	CGCTCMCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGAGCGGCGAGCAGA 811
QY	7279	AGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAGAGCTAGA 7338
Db	810	AGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAGAGCTAGA 751
QY	7339	GTAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTTGTGCAATTTGTTACAGGCATCGT 7398
Db	750	GTAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTTGTGCAATTTGTTACAGGCATCGT 691
QY	7399	GTGTACAGCTCGTGTGGTATGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGCA 7458
Db	690	GTGTACAGCTCGTGTGGTATGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGCA 631
QY	7459	GTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTT 7518
Db	630	GTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTT 572
QY	7519	GTCAG-AGTAAAGTTGCCCGCAGTGTATATCTCATCGTTATGCGACAGCTGCATAATTC 7577



Qy	7940	TTTTTCAATATATTGAAGCATTTATTCATCATGAGCGGATACATATTG	7999
Db	170	TTTTTCAATATATTGAAGCATTTATTCATCATGAGCGGATACATATTG	111
Qy	8000	AAATGATTTAGAAAAATA	8017
Db	110	AAATGATTTAGAAAAATA	93

RESULT 13				
CD649375/c				
LOCUS	CD649375	819 bp	mrna	linear
DEFINITION	CyGnd00080	Crassostrea virginica	Gonad	Crassostrea virginica cDNA
				5', mRNA sequence.

ACCESSION CD649375  
VERSION CD649375.1  
KEYWORDS GI:31906346  
EST.

SOURCE	ORGANISM
221	<i>Crassostrea virginica</i> (eastern oyster)
222	<i>Crassostrea virginica</i>
223	Eukaryota; Metazoa; Mollusca; Bivalvia; Pelecymorpha; Ostreoida; Ostreidae; Ostreidae; Crassostrea.

REFERENCE  
1 (Pages 1 to 819)  
AUTHORS  
Peatman,E., Kucuktas,H., Li,P., He,C., Feng,J., Wei,X. and Liu,Z.  
TITLE  
Differentially expressed cyster (*Crassostrea virginica*) genes after  
exposure to mercury  
JOURNAL  
Unpublished (2003)

CONTACT: Lin ZJ  
The Fish Molecular Genetics and Biotechnology Laboratory,  
Department of Fisheries and Allied Aquacultures and Program of Cell  
and Molecular Biosciences  
Auburn University  
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA  
Tel: 334 844 4054  
Fax: 334 844 9208  
Email: zliu@acesag.auburn.edu

```

FEATURES
    source
        Seq primer: M13 Reverse.
        Location/Qualifiers
            1..819
                /organism="Crassostrea virginica"
                /mol_type="mRNA"
                /db_xref="taxon:6565"
                /clone_lib="Crassostrea virginica Gonad"
                /note="Organ: Gonad; Vector: pSport1; Site_1: NoI1;
                Site 2: SalI"

```

```

ORIGIN                               Date_: 2011-07-16
Query Match                          9.8%; Score 804.8; DB 14; Length 819;
Best Local Similarity                99.6%; Pred. No. 4.1e+168;
Matches 817; Conservative            0; Mismatches 2; Indels 1; Gaps 1;

```

Qy	7067	ACAGTTACCAATGCTTAATCAGTGAGSCACCTATCTCAGGCACTGTGTCTATTTTGGTTTCAT	7126
Db	819	ACAGTTACCAATGCTTAATCAGTGAGSCACCTATCTCAGGCACTGTGTCTATTTTGGTTTCAT	760

Qy 7127 CCATAGTTGGCTGACTCCCCGTCGTGTAGATACTACGATACGGAGGGCTTACCATCTG 7186

Db 759 CCATAGTTGCTGACTCCCCGTCGTGATAGTACGATACGGAGGGCTTACCATCTG 700

7187 G C C C C A G T G C C A A T G A T A C C G C G A G A C C C A C G C T A C C G G C T C C A G A T T A T C A G C A A 640  
 Db

QY 7247 TAACGAGCCAGCCGGAAGGGCGGAGCGAGAAGTGGTCTGCAACTTTATCGGCTCCA 7306

D6 TAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTCTGCACTTTATCCGCGCTCCA 580

D<sub>b</sub> 579 TCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTC 520

Qy	7367	GCAACGTTGTTGCCATTGCTGTCAGGACGTCGTGGTGTACGCTCGTCGTTTGGTATGGCTT	7426
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			

DD 519 GCAACGTTGTATGCCCATTGGCTACAGGCAATCGTGGTGTACGCTCGTCGTTTGGGATATGGCTT 460

Qy	7427	CATTTCAGCTCCGGTTCCCAACGATCAAGCGAGTTTCATGATCCCCCATGTTTGTGCAAAA	7486
Db	459	CATTTCAGCTCCGGTTCCCAACGATCAAGCGAGTTTCATGATCCCCCATGTTTGTGCAAAA	400
Qy	7487	AAGCGGTTAGCTCCTTCGGTCCCTCCGATCGTTCTCAGAAGTAAGTTGGCGCGCAGTGTAT	7546
Db	399	AAGCGGTTAGCT-CCTCGGTCCCTCCGATCGTTCTCAGAAGTAAGTTGGCGCGCAGTGTAT	341
Qy	7547	CATCATGTTATGGGAGCACTGCATAAATCTCTTACTGTCATGCCATCCGTAAGATGCT	7606
Db	340	CATCATGTTATGGGAGCACTGCATAAATCTCTTACTGTCATGCCATCCGTAAGATGCT	281
Qy	7607	TTTCTGTGACTGGTGTAGTACTCAACCAAGTCATTTCTGAGAATAGTGTATGCGCGCACCGA	7666
Db	280	TTTCTGTGACTGGTGTAGTACTCAACCAAGTCATTTCTGAGAATAGTGTATGCGCGCACCGA	221
Qy	7667	GTTTGTCTTTCCCGGGGTCAACACGGGATAATACCCGGCCACATACAGAACTTTTAAAG	7726
Db	220	GTTTGTCTTTCCCGGGGTCAANTACGGGATAATACCCGGCCACATACAGAACTTTTAAAG	161
Qy	7727	TGTCATCATTTGAAAAACGTTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGA	7786
Db	160	TGTCATCATTTGAAAAACGTTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGA	101
Qy	7787	GATCCAGTTCCGATGTAAACCACTCGTGCAACCACTGATCTTTCAGCATCTTTTACTTTCA	7846
Db	100	GATCCAGTTCCGATGTAAACCACTCGTGCAACCACTGATCTTTCAGCATCTTTTACTTTCA	41
Qy	7847	CCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGC	7886
Db	40	CCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGC	1

RESULT 14					
B2576726/c					
LOCUS					
DEFINITION	B2576726	1011 bp	DNA	linear	GSS 17-DEC-2002
	msh2_5071.y2	msh	<i>Pseudomonas aeruginosa</i>	genomic clone	msh2_5071,
				genomic survey	sequence.

B2575726  
 B2575726.1 GI:27211787  
 GSS.  
 Pseudomonas aeruginosa  
 Pseudomonas aeruginosa  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 Pseudomonadaceae; Pseudomonas.

1 (bases 1 to 1011)  
 Spencer D.H., Raymond C.K., Smith E.E., Sims E.E., Hastings M.,  
 Burns J.L., Kaul R. and Olsen M.V.  
 Whole-Genome-Sequence variation among multiple isolates of  
*Pseudomonas aeruginosa* library  
 J. Bacteriol. (2002) in press  
 Contact: Chris K. Raymond  
 Genome Center

University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2063216954  
Fax: 2066857244  
Email: [craymond@u.washington.edu](mailto:craymond@u.washington.edu)  
Class: shotgun.

FEATURES	source
Location/Qualifiers	
1. .1011	
/organism="Pseudomonas aeruginosa"	
/mol_type="genomic DNA"	
/strain="M5H"	
/db_xref="taxon:287"	
/clone_lib="msh2_5071"	
/note="Environmental isolate. Whole genomic shotgun library"	

ORIGIN	Query Match	Score	DB 28;	Length 1011;
Library:		9.6%		





```

QY 6823 GGCAAAACAAACCAACCGCTGTAGCGTGG-TTTTTTTGTTTGCAGCAGAGATTACGCG 6881
Db |||||
QY 631 GGCAAAACAAACCAACCGCTGTAGCGTGGTTTTTTTGTGCAAGCAGCAGATTACGCG 690
Db |||||
QY 6882 CAGAAAAAAGGATCTCAGAGAGATCCTTTGATCTTTTCTACGGGTCTCAGCTCAGTG 6941
Db |||||
QY 691 CAGAAAAAAGGATCTCAGAGAGATCCTTTGATC-TTCTACGGGTCTCAGCTCAGTG 749
Db |||||
QY 6942 GAACGAAACCTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAAGGATCTTCACCTA 7001
Db |||||
QY 750 GAACGAAACCTCAGCTTAAGGGA-TTGGTCATGAGATTATCAAAAAAGGATCTTCACCTA 808
Db |||||
QY 7002 GATCCTTTTAAATTAATAAGTAAATCAATCTAAAGTATATGAGTAAACTTG 7061
Db |||||
QY 809 GATCCTTTTAAATTAATAAGTAAATCAAG-TTTAAATAATCTAAAGTATATGAGTAAACTTG 867
Db |||||
QY 7062 GTCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTGTCTATTTCG 7121
Db |||||
QY 868 GTCTGACAGTACCAATGGCTAATCAGTGAAGGACCTATCTTACCGATCTGTCTATTTCG 927
Db |||||
QY 7122 TTCAATCATAGTTGCCTGACTCCCGGTCGTGTAGATAACTACGATAGGGAGGGCTTACC 7181
Db |||||
QY 928 -TCATCCATAGTGGCCGACTCCCGGGTAAAAACAAC-----AATCCGAGGGCTTAC 981
Db |||||
QY 7182 ATCTGGCCCGAGTGTCAATGATACCGGAGAGACCCAGCTCACCAGCTCCAGATTATC 7241
Db |||||
QY 982 ATTGGGCCCCAGTGTGAAGTATACCGGAGAGACCCAGCTCCCGG-----TCCGATTATC 1036
Db |||||
QY 7242 AGCAATAAAC 7251
Db |||||
QY 1037 AGCAATAACC 1046
Db |||||

```

Search completed: September 24, 2004, 09:58:30  
 Job time : 12509.4 secs

This Page Blank (copy)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 02:36:16 ; Search time 1874.44 Seconds  
(without alignments)  
18677.278 Million cell updates/sec

Title: US-10-030-390-3

Perfect score: 8241  
Sequence: 1 agatctaatccaaagacg.....aaataacaggtattattcg 8241

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N Geneseq 29Jan04:\*
- 1: Geneseqn1980s:\*
  - 2: Geneseqn1990s:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002s:\*
  - 7: Geneseqn2003as:\*
  - 8: Geneseqn2003bs:\*
  - 9: Geneseqn2003cs:\*
  - 10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8241	100.0	8241	5	Aaf29647 Plasmid p
2	7555.6	91.7	8598	2	Aav28852 pPIC9/ELF
3	7260	88.1	8974	2	Aat77819 Neisseria
4	6687.6	81.2	9408	4	Aaf87135 Lunasin c
5	6675.6	81.0	9407	4	Aaf83654 pPIC9K-lu
6	6582	79.9	9191	2	Aat77818 Neisseria
7	6561.6	79.6	9276	4	Aad05811 Pichia pa
8	6561.6	79.6	9276	4	Aad13151 Pichia pa
9	6561.6	79.6	9276	7	Abz22459 Pichia pa
10	6561.6	79.6	9276	7	Ala60801 Pichia pa
11	6561.6	79.6	9276	7	Abz47231 pPIC9K ve
12	6561.6	79.6	9276	7	Aad47187 Pichia ve
13	6561.6	79.6	9276	7	Abz5813 Pichia pa
14	6561.6	79.6	9276	9	Abd97562 Pichia pa
15	6383.8	77.5	9776	6	Aad09083 Pichia pa
16	6383.8	77.5	9776	6	Aak99696 DNA of pl
17	6009	72.9	10215	2	Aat77820 Neisseria
18	5709.8	69.3	10462	4	Aad09086 pUB116 pl
19	5709.8	69.3	10462	6	Aak99699 DNA of pl
20	5677.8	68.9	10492	4	Aad09084 pUB114 pl
21	5665.8	68.8	10491	6	Aak99697 DNA of pl
22	5657.8	68.7	10512	4	Aad09085 pUB115 pl
23	5657.8	68.7	10512	6	Aak99698 DNA of pl

24	5142.8	62.4	8584	2	AAT35166	Aat35166 Plasmid p
25	5124.8	62.2	8590	2	AAT35168	Aat35168 Plasmid p
26	4734.6	57.5	9156	2	AAT77817	Aat77817 Neisseria
27	4555.2	55.4	8137	2	AAT35165	Aat35165 Plasmid p
28	2657.4	32.2	2659	2	AAT63530	Aat63530 Pichia pa
29	2404.6	29.2	5740	7	ACC42391	Acc42391 Plasmid p
30	2312.4	28.1	11466	6	ABQ74937	Abq74937 Vector py
31	2312.4	28.1	11466	6	ABK90548	Abk90548 Vector DN
32	2312.4	28.1	11466	6	AAK45828	Aak45828 pYAC4-Asc
33	2312.4	28.1	11466	6	ABK90626	Abk90626 DNA seque
34	2312.4	28.1	11466	9	ADC37166	Adc37166 Vector py
35	2311.2	28.0	11346	2	AAV12373	Aav12373 ChimERIC
36	2309.4	28.0	19798	4	AAF83671	Aaf83671 Nucleic a
37	2309.4	28.0	19912	4	AAF83670	Aaf83670 Nucleic a
38	2309.4	28.0	20160	4	AAF83672	Aaf83672 Nucleic a
39	2309.4	28.0	20217	4	AAF83674	Aaf83674 Nucleic a
40	2309.4	28.0	20247	4	AAF83675	Aaf83675 Nucleic a
41	2309.4	28.0	20316	4	AAF83673	Aaf83673 Nucleic a
42	2309.2	28.0	10111	6	AAH99982	Aah99982 pME908 re
43	2306.8	28.0	3621	1	AAAN60847	Aan60847 Human pre
44	2306	28.0	34303	2	AAV07261	Aav07261 Adenovira
45	2305.2	28.0	3547	1	AAAN60846	Aan60846 Plasmid s

ALIGNMENTS

RESULT 1  
AAF29647  
ID AAF29647 standard; DNA; 8241 BP.  
XX  
AC AAF29647;  
XX  
DT 10-APR-2001 (first entry)  
XX  
DE Plasmid pPICmTFF1.  
XX  
KW Mouse TFF1; mTFF; trefoil protein; antiinflammatory; gastrointestinal;  
KW anti-ulcer; peptide therapy; gastrointestinal disease; acute colitis;  
KW Crohn's disease; ulcerative colitis; plasmid pPICmTFF1;  
KW recombinant vector; ds.  
XX  
OS Synthetic.  
XX  
PN WO200102570-A1.  
XX  
PD 11-JAN-2001.  
XX  
PF 05-JUL-2000; 2000WO-EP006343.  
XX  
PR 05-JUL-1999; 99EP-00870143.  
XX  
PA (VIAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
XX  
PI Hans WC, Steidler L, Remaut ER;  
XX  
DR WPI; 2001-138142/14.  
XX  
PT Recombinant Lactococcus lactis for delivering a trefoil peptide useful  
PT for treating acute or chronic gastrointestinal inflammatory diseases or  
PT disorders, e.g. acute or ulcerative colitis, acute flare-ups of Crohn's  
PT disease.  
XX  
PS Claim 18; Fig 1c; 44pp; English.  
XX  
CC The present sequence is a recombinant vector which may be used to  
CC generate a recombinant Lactococcus lactis capable of delivering a trefoil  
CC peptide in vivo. The recombinant microorganism is useful for  
CC manufacturing an agent for the delivery of a trefoil peptide to the  
CC gastrointestinal tract, and for treating gastric or intestinal diseases  
CC or disorders, or lesions caused by gastric or intestinal diseases or  
CC disorders. The microorganism may also be used for preparing medicament to  
CC be used for treating gastric and/or gastrointestinal diseases or

CC disorders, acute gastrointestinal inflammatory diseases (e.g., acute  
 CC colitis, acute flare-ups of Crohn's diseases, or ulcerative colitis), and  
 CC chronic and spontaneously recurring diseases of the gastrointestinal  
 CC tract comprising Crohn's disease (enteritis regionalis) and ulcerative  
 CC colitis (colitis ulcerosa). Disease states which can be treated by the  
 CC method or compositions comprising the recombinant microorganism or  
 CC trefoil peptides include disorders of and damage to the alimentary canal,  
 CC including the mouth, esophagus, stomach and large and small intestine,  
 CC as well as for the protection and treatment of tissues that lie outside  
 CC the alimentary canal  
 XX

SQ Sequence 8241 BP; 2193 A; 1963 C; 1870 G; 2215 T; 0 U; 0 Other;

Query Match 100.0%; Score 8241; DB 5; Length 8241;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 8241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGATCTAAATCCAAAGAGGAAAGGTTGAATGAACCCCTTTTGCCATCCGACATCCACAG 60  
 DB 1 AGATCTAAATCCAAAGAGGAAAGGTTGAATGAACCCCTTTTGCCATCCGACATCCACAG 60  
 QY 61 GTCCATTCTCACATAAAGTGGCAACAGCAAGAGGGGATACACTAGCAGACCGGT 120  
 DB 61 GTCCATTCTCACATAAAGTGGCAACAGCAAGAGGGGATACACTAGCAGACCGGT 120  
 QY 121 TGCARACGAGGACCTCCACTCTCTCTCTCAACACCCCTTTTGCCATCGAAACCC 180  
 DB 121 TGCARACGAGGACCTCCACTCTCTCTCTCAACACCCCTTTTGCCATCGAAACCC 180  
 QY 181 AGCCAGATTATTGGGCTTGAATGGAGCTCGCTCATTCOAATTCCTTCTATTAGGCTACTA 240  
 DB 181 AGCCAGATTATTGGGCTTGAATGGAGCTCGCTCATTCOAATTCCTTCTATTAGGCTACTA 240  
 QY 241 ACACCATGACTTTATTAGGCTCTCTATCTGCCCCCTGGGAGGTTTCATGTTTGT 300  
 DB 241 ACACCATGACTTTATTAGGCTCTCTATCTGCCCCCTGGGAGGTTTCATGTTTGT 300  
 QY 301 TTTCGGAATGCAACAGCTCCGATTAACCCGAAACATCACTCCAGATGAGGGCTTTCTG 360  
 DB 301 TTTCGGAATGCAACAGCTCCGATTAACCCGAAACATCACTCCAGATGAGGGCTTTCTG 360  
 QY 361 AGTGTGGGCTCAATAGTTTCATGTTCCCAATAGCCCAAACTGACAGTTTAAAGCCT 420  
 DB 361 AGTGTGGGCTCAATAGTTTCATGTTCCCAATAGCCCAAACTGACAGTTTAAAGCCT 420  
 QY 421 GTCTTGGAACCTTAATATGACAAAGCGTGATCTCATCAAGATGAACCTTAAGTTGGTTCC 480  
 DB 421 GTCTTGGAACCTTAATATGACAAAGCGTGATCTCATCAAGATGAACCTTAAGTTGGTTCC 480  
 QY 481 TTGAATGTACGGCCAGTTGGTCAAAAGAACTTCCAAAGTCCGCATACCGTTTGT 540  
 DB 481 TTGAATGTACGGCCAGTTGGTCAAAAGAACTTCCAAAGTCCGCATACCGTTTGT 540  
 QY 541 CTGTGTTGATTTGATTTGACGAATGCTCAAAATAATCTCATTAATGCTTAGCGCAGTCT 600  
 DB 541 CTGTGTTGATTTGATTTGACGAATGCTCAAAATAATCTCATTAATGCTTAGCGCAGTCT 600  
 QY 601 CTCTATCGCTTCTGACCCCGTGACCTGTGCGGAACGCAATGGGAAACACCCGCT 660  
 DB 601 CTCTATCGCTTCTGACCCCGTGACCTGTGCGGAACGCAATGGGAAACACCCGCT 660  
 QY 661 TTTTGGATGATTATGCAATGTCTCCACATTTGATGCTTCCCAAGATTCTGGTGGGAATACT 720  
 DB 661 TTTTGGATGATTATGCAATGTCTCCACATTTGATGCTTCCCAAGATTCTGGTGGGAATACT 720  
 QY 721 GCTGTAGCTTAACCTGTCAATGCAAAATTTAACTGTTCTAACCCCTACTTGCACGAAT 780  
 DB 721 GCTGTAGCTTAACCTGTCAATGCAAAATTTAACTGTTCTAACCCCTACTTGCACGAAT 780  
 QY 781 ATATAACAGAGAGGAGTGGCTCTCTTAAACCTTTTCTTAACTCATCATTTAGCTT 840  
 DB 781 ATATAACAGAGAGGAGTGGCTCTCTTAAACCTTTTCTTAACTCATCATTTAGCTT 840

QY 841 ACTTTTCATAATTGCGACTGGTTCGAATTGACAAAGCTTTTGATTTTAAACGACTTTTAAACGA 900  
 DB 841 ACTTTTCATAATTGCGACTGGTTCGAATTGACAAAGCTTTTGATTTTAAACGACTTTTAAACGA 900  
 QY 901 CAACTTGAGAAGATCAAAAAACAACTAATTTATTCGAAGGATCCAAACGATGAGATTTCCT 960  
 DB 901 CAACTTGAGAAGATCAAAAAACAACTAATTTATTCGAAGGATCCAAACGATGAGATTTCCT 960  
 QY 961 TCAATTTTACTGCGAGTTTATTCGCGACTCTCCGCAATTCCTCCGATAGCTCTCCAGTCAACACT 1020  
 DB 961 TCAATTTTACTGCGAGTTTATTCGCGACTCTCCGCAATTCCTCCGATAGCTCTCCAGTCAACACT 1020  
 QY 1021 ACAACAGAAGATGAAACCGCACAAATTCGGCTGAAAGCTGTCTATCGGTTACTCAGATTTA 1080  
 DB 1021 ACAACAGAAGATGAAACCGCACAAATTCGGCTGAAAGCTGTCTATCGGTTACTCAGATTTA 1080  
 QY 1081 GAAGGGGATTTGCGATGTTGCTGTTTCCCATTTTCCAAACGACCAATTAACGGGTTATTG 1140  
 DB 1081 GAAGGGGATTTGCGATGTTGCTGTTTCCCATTTTCCAAACGACCAATTAACGGGTTATTG 1140  
 QY 1141 TTTATAAATACTACTATTGCGACATTCGCTCTCTAAAGAAGAGGGGTATCTCTCGAGAA 1200  
 DB 1141 TTTATAAATACTACTATTGCGACATTCGCTCTCTAAAGAAGAGGGGTATCTCTCGAGAA 1200  
 QY 1201 AGAGAGGCTGAAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCAAGAAACATCTATCATG 1260  
 DB 1201 AGAGAGGCTGAAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCAAGAAACATCTATCATG 1260  
 QY 1261 GCCCCCCGGGAGAGGATAAATTTGGCTTCCCGGTTCACCCGCCAGCAGTGCACGGAG 1320  
 DB 1261 GCCCCCCGGGAGAGGATAAATTTGGCTTCCCGGTTCACCCGCCAGCAGTGCACGGAG 1320  
 QY 1321 AGAGGTTCTGTTTGTATGACAGTCCGGGATCCCGTGGTCTCCACCCCGAGTGCACGGCC 1380  
 DB 1321 AGAGGTTCTGTTTGTATGACAGTCCGGGATCCCGTGGTCTCCACCCCGAGTGCACGGCC 1380  
 QY 1381 ATCGAGAACACTCAAGAGAGAAATGTCCCTTTCTAACTAGTGGCTAGAAATTCCTTAGGG 1440  
 DB 1381 ATCGAGAACACTCAAGAGAGAAATGTCCCTTTCTAACTAGTGGCTAGAAATTCCTTAGGG 1440  
 QY 1441 CGGCGCGCAATTAATTCGCTTAGACACTGTTCTCTCAGTTCAAGTTGGGCACTTAG 1500  
 DB 1441 CGGCGCGCAATTAATTCGCTTAGACACTGTTCTCTCAGTTCAAGTTGGGCACTTAG 1500  
 QY 1501 AGAAGACCGGCTTCTGCTAGATTCTTAATCAAGAGGATGTCAAGATGCCATTTCGCTCAGAG 1560  
 DB 1501 AGAAGACCGGCTTCTGCTAGATTCTTAATCAAGAGGATGTCAAGATGCCATTTCGCTCAGAG 1560  
 QY 1561 ATGCAAGGCTTCATTTTGTACTTTTATTTTAACTATATAGTATAGGATTTTTTTTT 1620  
 DB 1561 ATGCAAGGCTTCATTTTGTACTTTTATTTTAACTATATAGTATAGGATTTTTTTTT 1620  
 QY 1621 GTCATTTTGTCTTCTCTGTAACGAGCTTGTCTCTGATCAGCCTATCTCGCAGCTGATGAA 1680  
 DB 1621 GTCATTTTGTCTTCTCTGTAACGAGCTTGTCTCTGATCAGCCTATCTCGCAGCTGATGAA 1680  
 QY 1681 TATCTGTTGGTAGGGTTTGGGAAATTCATTCAGATTGATGTTTCTTGGTATTTCCT 1740  
 DB 1681 TATCTGTTGGTAGGGTTTGGGAAATTCATTCAGATTGATGTTTCTTGGTATTTCCT 1740  
 QY 1741 ACTCCTCTCAGAGTACAGAAGATTAAGTCAGAAGTTGCTTGTGCAAGCTTATCGATAA 1800  
 DB 1741 ACTCCTCTCAGAGTACAGAAGATTAAGTCAGAAGTTGCTTGTGCAAGCTTATCGATAA 1800  
 QY 1801 GCTTTAATGCGGTATTTATCAGATTTAATTCCTAACCGAGTCAGCCCGGTGATGAA 1860  
 DB 1801 GCTTTAATGCGGTATTTATCAGATTTAATTCCTAACCGAGTCAGCCCGGTGATGAA 1860  
 QY 1861 ATCTAAACAATGCGCTCATCTCATCTCCGCAACCGCTCACCCCTGATGCTGTAGGCAATAG 1920  
 DB 1861 ATCTAAACAATGCGCTCATCTCATCTCCGCAACCGCTCACCCCTGATGCTGTAGGCAATAG 1920  
 QY 1921 CTTGGTTATCCGGTACTGCGCGGCGCTCTTTCGGGGATATCGTCCATTCGACAGCATCGC 1980

Db 1921 CTTGGTATGCGGTACTGCCGGCCCTCTTGGGGATATCGTCCATTCGGACAGATCGC 1980  
Qy 1981 CAGTCATATGCGGTGCTGTAGCGCTATATGCGTTGATGCAATTCCTATGCGCACCGGT 2040  
Db 1981 CAGTCATATGCGGTGCTGTAGCGCTATATGCGTTGATGCAATTCCTATGCGCACCGGT 2040  
Qy 2041 TCTCGGAGCACTGTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGTTCGCTACTTGG 2100  
Db 2041 TCTCGGAGCACTGTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGTTCGCTACTTGG 2100  
Qy 2101 AGCCACTATGCACTACCGCATCATGGCGACACACCCCGTCTGTGGATCTATCGAATCTA 2160  
Db 2101 AGCCACTATGCACTACCGCATCATGGCGACACACCCCGTCTGTGGATCTATCGAATCTA 2160  
Qy 2161 AATGTAAGTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220  
Db 2161 AATGTAAGTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220  
Qy 2221 CAGCATTCGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA 2280  
Db 2221 CAGCATTCGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA 2280  
Qy 2281 TATGTTTCAGTCCCTCAGAGTTAGCTTCTGGAAGTATGAACTTCTGGAAGTTCGAG 2340  
Db 2281 TATGTTTCAGTCCCTCAGAGTTAGCTTCTGGAAGTATGAACTTCTGGAAGTTCGAG 2340  
Qy 2341 TGTAACTCCGCTGATTAAGCGGCAATCCGTCAGTTGGCAAGTGGTGGTACCGG 2400  
Db 2341 TGTAACTCCGCTGATTAAGCGGCAATCCGTCAGTTGGCAAGTGGTGGTACCGG 2400  
Qy 2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGCAACAAACAAACAGATCCAGCGTGT 2460  
Db 2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGCAACAAACAAACAGATCCAGCGTGT 2460  
Qy 2461 GTACTTGATCAACATAGAAGAGCATCTCGATTTGAGGATCAAGTGTTCAGAGCGT 2520  
Db 2461 GTACTTGATCAACATAGAAGAGCATCTCGATTTGAGGATCAAGTGTTCAGAGCGT 2520  
Qy 2521 ACTGATTTGGACATTTCAAAGCTGCTGTAGTTGCAACCGATAGGTTGTAGAGTGTG 2580  
Db 2521 ACTGATTTGGACATTTCAAAGCTGCTGTAGTTGCAACCGATAGGTTGTAGAGTGTG 2580  
Qy 2581 CAATACACTTGGTACATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAACAGCAT 2640  
Db 2581 CAATACACTTGGTACATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAACAGCAT 2640  
Qy 2641 CTTCAATTTCTGGCAAGCTTCTGTCTATATCGACAGCCAAACAGATCACTTGGGAAT 2700  
Db 2641 CTTCAATTTCTGGCAAGCTTCTGTCTATATCGACAGCCAAACAGATCACTTGGGAAT 2700  
Qy 2701 CAATACCATTTGCTGCTGAGACAGAGTCTGAGGCAACGAATCTGGATCAGCGTATT 2760  
Db 2701 CAATACCATTTGCTGCTGAGACAGAGTCTGAGGCAACGAATCTGGATCAGCGTATT 2760  
Qy 2761 TATCAGCAATTAACATAGAACTTCAAGAGCCAGAGGCAACGAATCTGGATCAGCGTATT 2820  
Db 2761 TATCAGCAATTAACATAGAACTTCAAGAGCCAGAGGCAACGAATCTGGATCAGCGTATT 2820  
Qy 2821 ATGTCATTTTGAACCATCTATTTGGCAGCAGTAAACGAATCTGGATCAGCGTATT 2880  
Db 2821 ATGTCATTTTGAACCATCTATTTGGCAGCAGTAAACGAATCTGGATCAGCGTATT 2880  
Qy 2881 TTTTGTACACTTAGGAACAGTTCTGTTCCGTAGCCATAGCAGTACTGCTGGGCGC 2940  
Db 2881 TTTTGTACACTTAGGAACAGTTCTGTTCCGTAGCCATAGCAGTACTGCTGGGCGC 2940  
Qy 2941 CTCCTGCTAGCAGATACACTTAGCAACCTTGTGGGCAACGATAGTACTTCTGGG 3000  
Db 2941 CTCCTGCTAGCAGATACACTTAGCAACCTTGTGGGCAACGATAGTACTTCTGGG 3000  
Qy 3001 TAAGGTAACCATCTTCTTAGGTGAGTAGTGAACCAATTTCTTTTCCAAACCGACATCTT 3060

Db 3001 TAAGGTAACCATCTTCTTAGGTGAGTAGTGAACCAATTTCTTTTCCAAACCGACATCTT 3060  
Qy 3061 TGGCAGAAACACCCAGCATCAGGAAGTGGAAAGCGAGAAATTCGGTTCACACAGGAATAT 3120  
Db 3061 TGGCAGAAACACCCAGCATCAGGAAGTGGAAAGCGAGAAATTCGGTTCACACAGGAATAT 3120  
Qy 3121 AGAGGCCAACTTTCTCAATAGTCTTTCGCAAAACGAGAGCAGACTACACAGAGGCAAGTCT 3180  
Db 3121 AGAGGCCAACTTTCTCAATAGTCTTTCGCAAAACGAGAGCAGACTACACAGAGGCAAGTCT 3180  
Qy 3181 CAACTTTCGAAACCGTCTCCGTTAGTTCATGGAATTTCTTCGAGCTTATCTATAGAGA 3240  
Db 3181 CAACTTTCGAAACCGTCTCCGTTAGTTCATGGAATTTCTTCGAGCTTATCTATAGAGA 3240  
Qy 3241 GATCAATGCTCTCTTAACGTTATCTGCAATTCGCAATTCCTCTGCGGAAAGGAGCTT 3300  
Db 3241 GATCAATGCTCTCTTAACGTTATCTGCAATTCGCAATTCCTCTGCGGAAAGGAGCTT 3300  
Qy 3301 CTAACACAGGTGTCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTCTTAAAGGGCTT 3360  
Db 3301 CTAACACAGGTGTCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTCTTAAAGGGCTT 3360  
Qy 3361 TGTCAACATTTTGAAGCAATTTGCGCAATTTGCTGAAATTCCTAATTCCTGTCG 3420  
Db 3361 TGTCAACATTTTGAAGCAATTTGCGCAATTTGCTGAAATTCCTAATTCCTGTCG 3420  
Qy 3421 TTTTCTGATAGCAGCAAGCGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAACGT 3480  
Db 3421 TTTTCTGATAGCAGCAAGCGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAACGT 3480  
Qy 3481 CAATTTGCACAATTTCAATACGACTTCAAGAGGAGCTTCTTTAGTTTGGATCTTCTT 3540  
Db 3481 CAATTTGCACAATTTCAATACGACTTCAAGAGGAGCTTCTTTAGTTTGGATCTTCTT 3540  
Qy 3541 TAGTTTGTCTTCTTGGTGTATCTTGGCTTGGCATCTCTTTCTCTTCTAGTACCTTTAGG 3600  
Db 3541 TAGTTTGTCTTCTTGGTGTATCTTGGCTTGGCATCTCTTTCTCTTCTAGTACCTTTAGG 3600  
Qy 3601 ACTTCATATCCAGTTTCTCTCCACTCGTCCAAAGTCAACCGTCAACCGTCACTTGGACATCTAA 3660  
Db 3601 ACTTCATATCCAGTTTCTCTCCACTCGTCCAAAGTCAACCGTCAACCGTCACTTGGACATCTAA 3660  
Qy 3661 CTAATCCAAATTAATAATAGTCAGCATCTCCAGGCTATATCTTCTTGGATTTAGCTT 3720  
Db 3661 CTAATCCAAATTAATAATAGTCAGCATCTCCAGGCTATATCTTCTTGGATTTAGCTT 3720  
Qy 3721 CTGCAAGTTTCATCAGCTTCTCTCCCTAATTTAGCGTTCAACAAACTTCTCGTCAAAATA 3780  
Db 3721 CTGCAAGTTTCATCAGCTTCTCTCCCTAATTTAGCGTTCAACAAACTTCTCGTCAAAATA 3780  
Qy 3781 ACGTTTGGTATAAGAACCTTCTGAGCATTCCTTCTAGCATCCCAAGGTGGCTTCCA 3840  
Db 3781 ACGTTTGGTATAAGAACCTTCTGAGCATTCCTTCTAGCATCCCAAGGTGGCTTCCA 3840  
Qy 3841 TGGCTCTAAGACCTTTGATTTGGCCAAACAGAGAGTGGGTTCCAGAGTGAAGAACCAA 3900  
Db 3841 TGGCTCTAAGACCTTTGATTTGGCCAAACAGAGAGTGGGTTCCAGAGTGAAGAACCAA 3900  
Qy 3901 CACCTGTTTGTTCACCAACCAATTTCAAGAGTTCCTATCAAACTTCCATCCCAATTCGATACCA 3960  
Db 3901 CACCTGTTTGTTCACCAACCAATTTCAAGAGTTCCTATCAAACTTCCATCCCAATTCGATACCA 3960  
Qy 3961 GCAACTTTTGTGTTGCTCCAGATGAGCACCTTTTATACCAAAACCGTGACGACGAGATT 4020  
Db 3961 GCAACTTTTGTGTTGCTCCAGATGAGCACCTTTTATACCAAAACCGTGACGACGAGATT 4020  
Qy 4021 GGTAGACTCCAGTTTGTGCTTATAGCTTCGGAAATAGACTTTTGGACGAGTACACCA 4080  
Db 4021 GGTAGACTCCAGTTTGTGCTTATAGCTTCGGAAATAGACTTTTGGACGAGTACACCA 4080  
Qy 4081 GGGCCCAACGAGTAATTAGAAGAGTCAAGCACCAAAAGTAGTGAATAGACCATCGGGCGGT 4140  
Db 4081 GGGCCCAACGAGTAATTAGAAGAGTCAAGCACCAAAAGTAGTGAATAGACCATCGGGCGGT 4140

4141 CAGTAGTCAAGAGCGCAACAAAATTTCACTGACAGGAACTTTTGTACATCTTCAGAAA 4200  
 Db CAGTAGTCAAGAGCGCAACAAAATTTCACTGACAGGAACTTTTGTACATCTTCAGAAA 4200  
 4201 GTTCGTAATCAGTAGTCAATTCGCGAGCATCAATATATGGGATATATACAGAGCAACAG 4260  
 Db GTTCGTAATCAGTAGTCAATTCGCGAGCATCAATATATGGGATATATACAGAGCAACAG 4260  
 4261 TGGAGTCAACATCTACCAACTTTTGGCGTCTCAGAAAAGCATAAAGTTCTTACTACCGC 4320  
 Db TGGAGTCAACATCTACCAACTTTTGGCGTCTCAGAAAAGCATAAAGTTCTTACTACCGC 4320  
 4321 CATATGAACTTTTCAATTCGCCAGTGGGAGAGAAAGGCAAGCGATCTAGCAT 4380  
 Db CATATGAACTTTTCAATTCGCCAGTGGGAGAGAAAGGCAAGCGATCTAGCAT 4380  
 4381 TAGCGGCAAGGATGCAACTTTTCAACAGGCTTCTATAGATAACCTAGCGCTGGGA 4440  
 Db TAGCGGCAAGGATGCAACTTTTCAACAGGCTTCTATAGATAACCTAGCGCTGGGA 4440  
 4441 TCATCCCTTTGGCAACTTTTCTGCCAAATCTAGGTCGAAATCACTTCAATGATACCAT 4500  
 Db TCATCCCTTTGGCAACTTTTCTGCCAAATCTAGGTCGAAATCACTTCAATGATACCAT 4500  
 4501 TATTGTACAACTTACGCAAGTTGCGATCAGCTCCTCAAAATGCTCTGTAAACGATG 4560  
 Db TATTGTACAACTTACGCAAGTTGCGATCAGCTCCTCAAAATGCTCTGTAAACGATG 4560  
 4561 ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATGAGTGAATCTGATCAGGTTGTGCA 4620  
 Db ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATGAGTGAATCTGATCAGGTTGTGCA 4620  
 4621 GCTGGTCAGAGCATAGGGAACACGGCTTTTCTACCAACTCAAGGAATATCAAACT 4680  
 Db GCTGGTCAGAGCATAGGGAACACGGCTTTTCTACCAACTCAAGGAATATCAAACT 4680  
 4681 CTGCAACACTTGCATGAGTGAAGGAAATGTCATCTGAACTCGGACAGTGAG 4740  
 Db CTGCAACACTTGCATGAGTGAAGGAAATGTCATCTGAACTCGGACAGTGAG 4740  
 4741 TGTAGTCTTGAGAAATCTGAAGCGCTATTTTATATATCAGTAGTCACTCAGAGAG 4800  
 Db TGTAGTCTTGAGAAATCTGAAGCGCTATTTTATATATCAGTAGTCACTCAGAGAG 4800  
 4801 TCCTCTAGCGGACGATCTGCGGACCTGCAAGTCCGATCACCAGCGCCACAGGTG 4860  
 Db TCCTCTAGCGGACGATCTGCGGACCTGCAAGTCCGATCACCAGCGCCACAGGTG 4860  
 4861 CGGTTGCTGGGCGCTATATCCCGACATCAACGATGGGGAAGATCGGGCTCGCCACTTCG 4920  
 Db CGGTTGCTGGGCGCTATATCCCGACATCAACGATGGGGAAGATCGGGCTCGCCACTTCG 4920  
 4921 GCTCATGAGCGCTTTTTCGGCTGGGATAGTGGCAGGCCCCCGTGGCGGGGACTGT 4980  
 Db GCTCATGAGCGCTTTTTCGGCTGGGATAGTGGCAGGCCCCCGTGGCGGGGACTGT 4980  
 4981 TGGGCGCATCTCTTGCATGCAACATCTTTCGCGGCGGCTGCTCAACGGGCTCAACC 5040  
 Db TGGGCGCATCTCTTGCATGCAACATCTTTCGCGGCGGCTGCTCAACGGGCTCAACC 5040  
 5041 TACTACTGGGCTGCTTCTTAATGAGGATGCGATAGGAGAGCGTGGATCTATGA 5100  
 Db TACTACTGGGCTGCTTCTTAATGAGGATGCGATAGGAGAGCGTGGATCTATGA 5100  
 5101 TTGGAAGTATGGGAATGGTATACCGCATTTCTTCACTGCTCTTCACTATCAGAT 5160  
 Db TTGGAAGTATGGGAATGGTATACCGCATTTCTTCACTGCTCTTCACTATCAGAT 5160  
 5161 TATGCCCACTAAAGCAACCGGAGGAGATTTTCACTGGTAAATTTCTTCACTTTTGGT 5220  
 Db TATGCCCACTAAAGCAACCGGAGGAGATTTTCACTGGTAAATTTCTTCACTTTTGGT 5220

5221 CATCAGTAGACTGAACTGTGAGACTATCTCGGTTATGACAGCAGAAAATCTCTCTCTGG 5280  
 Db CATCAGTAGACTGAACTGTGAGACTATCTCGGTTATGACAGCAGAAAATCTCTCTCTGG 5280  
 5281 AGACAGTAAATGAAGTCCCAACCAATAAAGAAATCTTGTATCAGGAACAACTCTCTGT 5340  
 Db AGACAGTAAATGAAGTCCCAACCAATAAAGAAATCTTGTATCAGGAACAACTCTCTGT 5340  
 5341 TTCGAACTTTTTCGGTCTGCTGAACTATAAATCTAGTGGATATCTCGGTTAGGAATG 5400  
 Db TTCGAACTTTTTCGGTCTGCTGAACTATAAATCTAGTGGATATCTCGGTTAGGAATG 5400  
 5401 GAGCGGCAAAATGCTTACCTTCTGCACTTCAAGAGGTATGTAGGTTCTAGATACTGA 5460  
 Db GAGCGGCAAAATGCTTACCTTCTGCACTTCAAGAGGTATGTAGGTTCTAGATACTGA 5460  
 5461 TGCCAACTTCAAGTCAACAGTTCCTTTCGTTCAACCACTTCGATCCGATCCAGGAATCA 5520  
 Db TGCCAACTTCAAGTCAACAGTTCCTTTCGTTCAACCACTTCGATCCGATCCAGGAATCA 5520  
 5521 AAGTTGTTGCTACTATTCGATCCAGCCAGTGGGCTCTTGAACCTGCAATAGTGTCT 5580  
 Db AAGTTGTTGCTACTATTCGATCCAGCCAGTGGGCTCTTGAACCTGCAATAGTGTCT 5580  
 5581 CGTGTGTTGAGGTCACTCTTGTATGAATAAATCTAGTCTTGTATCTAAATACTTGACG 5640  
 Db CGTGTGTTGAGGTCACTCTTGTATGAATAAATCTAGTCTTGTATCTAAATACTTGACG 5640  
 5641 AGCCNAGCGTAATAACCCAAATCTAAATCTTTTAAACCTTTTAAAGGACAAAGTATG 5700  
 Db AGCCNAGCGTAATAACCCAAATCTAAATCTTTTAAACCTTTTAAAGGACAAAGTATG 5700  
 5701 TCTGCTGTATTAACCCAAATCAGCTCGTAGTCTGATCTCTCACTCAACTTGAGGGCAC 5760  
 Db TCTGCTGTATTAACCCAAATCAGCTCGTAGTCTGATCTCTCACTCAACTTGAGGGCAC 5760  
 5761 TATCTGTTTAGAGAAATTTTCGGAGATCGATCGAGAAAGGTAGCTGATTTTA 5820  
 Db TATCTGTTTAGAGAAATTTTCGGAGATCGATCGAGAAAGGTAGCTGATTTTA 5820  
 5821 AACGTAAATTTATCTCAAGATCTCTGCTCGCGCTTTCGGTGTATGACGCTGAAACCT 5880  
 Db AACGTAAATTTATCTCAAGATCTCTGCTCGCGCTTTCGGTGTATGACGCTGAAACCT 5880  
 5881 CTGACATATGAGCTCCCGAGACGGTCAAGCTTGTCTGTAAGCGGATCCGGAGCAG 5940  
 Db CTGACATATGAGCTCCCGAGACGGTCAAGCTTGTCTGTAAGCGGATCCGGAGCAG 5940  
 5941 ACAAGCCGCTCAGGCGGCTCAGCGGCTTTCGGGCTGTCGGGCGCAGCCATGACCCA 6000  
 Db ACAAGCCGCTCAGGCGGCTCAGCGGCTTTCGGGCTGTCGGGCGCAGCCATGACCCA 6000  
 6001 GTACGTAGCGATAGCGGAGTGTATCTGCTTAACTATGCGGATCAGAGCAGATTTGA 6060  
 Db GTACGTAGCGATAGCGGAGTGTATCTGCTTAACTATGCGGATCAGAGCAGATTTGA 6060  
 6061 CTGAGAGTGCACCATATCGGTTGTAATACCGACAGATCGTAAAGGAAATACCGC 6120  
 Db CTGAGAGTGCACCATATCGGTTGTAATACCGACAGATCGTAAAGGAAATACCGC 6120  
 6121 ATCAGCGCTCTTTCGGCTTCTCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 6180  
 Db ATCAGCGCTCTTTCGGCTTCTCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 6180  
 6181 CGAGCGGTATCAGCTCACTCAAGCGGTATACGTTATCCACAGATCAGGGATTAAC 6240  
 Db CGAGCGGTATCAGCTCACTCAAGCGGTATACGTTATCCACAGATCAGGGATTAAC 6240  
 6241 GCAGGAAGAAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGGCGCG 6300  
 Db GCAGGAAGAAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGGCGCG 6300  
 6301 TTGCTGGCGTTTTCATAGGCTCGCCCCCTGACGAGCATCAAAAATCGAGCTCA 6360

6301 TTGCTGGGCTTTTTCATAGGCTCGCCCTCGAGGACATCAAAAATCGACGCTCA 6360  
6361 AGTCAGAGTGGCGAAACCCGACGAGACTATAAGATACCGAGCGTTTTCGCCCTGGAGC 6420  
6361 AGTCAGAGTGGCGAAACCCGACGAGACTATAAGATACCGAGCGTTTTCGCCCTGGAGC 6420  
6421 TCCCTCGTGGCTCTCCTGTTCCGACCTCGCCCTTACCGGATACCTGTCCGCTTCTC 6480  
6421 TCCCTCGTGGCTCTCCTGTTCCGACCTCGCCCTTACCGGATACCTGTCCGCTTCTC 6480  
6481 CCTTCGGGAGAGTGGCGCTTCTCATAGCTACAGCTAGGTATCTCAGTTCCGTTAG 6540  
6481 CCTTCGGGAGAGTGGCGCTTCTCATAGCTACAGCTAGGTATCTCAGTTCCGTTAG 6540  
6541 GTGCTTCGCTCAAGCTGGGCTGTGCAAGAACCCCGCTTACCGCCGACCGCTCGCC 6600  
6541 GTGCTTCGCTCAAGCTGGGCTGTGCAAGAACCCCGCTTACCGCCGACCGCTCGCC 6600  
6601 TTATCCGGTAACTATCGTCTTGAAGTCCAAACCGGTAAGACAGCTTATCGCCACTGGCA 6660  
6601 TTATCCGGTAACTATCGTCTTGAAGTCCAAACCGGTAAGACAGCTTATCGCCACTGGCA 6660  
6661 CGAGCACTGGTAACAGGATAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTG 6720  
6661 CGAGCACTGGTAACAGGATAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTG 6720  
6721 AAGTGGTGGCTTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTGGCTCTGCTG 6780  
6721 AAGTGGTGGCTTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTGGCTCTGCTG 6780  
6781 AAGCAGTACCTTCGGGAAAGAGTGGTGGTGTGATCGGCAACAAACACCGCT 6840  
6781 AAGCAGTACCTTCGGGAAAGAGTGGTGGTGTGATCGGCAACAAACACCGCT 6840  
6841 GGTAGCGTGGTGTGTTTGTGAAAGCAGAGATACCGCAGAGGAGGATCTCAA 6900  
6841 GGTAGCGTGGTGTGTTTGTGAAAGCAGAGATACCGCAGAGGAGGATCTCAA 6900  
6901 GAAGATCTTTGATCTTTTACGGGCTGACGCTGAGTGAAGAGGAACTACGTTAA 6960  
6901 GAAGATCTTTGATCTTTTACGGGCTGACGCTGAGTGAAGAGGAACTACGTTAA 6960  
6961 GGGATTTGGTCAATGAGATATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATA 7020  
6961 GGGATTTGGTCAATGAGATATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATA 7020  
7021 TGAAGTTTAAATCAATCTAAAGTATATAGTAAGTCTGGTGTGACAGTTACCAATGC 7080  
7021 TGAAGTTTAAATCAATCTAAAGTATATAGTAAGTCTGGTGTGACAGTTACCAATGC 7080  
7081 TTAATCAGTGGGACCTATCTCAGGATCTGTCTATTTCTGATCATAGTTGCTGTA 7140  
7081 TTAATCAGTGGGACCTATCTCAGGATCTGTCTATTTCTGATCATAGTTGCTGTA 7140  
7141 CTCCCGCTCGTGTAGATAACTACGATACGAGGAGGCTTACCATCTGGCCCGAGTGTGCA 7200  
7141 CTCCCGCTCGTGTAGATAACTACGATACGAGGAGGCTTACCATCTGGCCCGAGTGTGCA 7200  
7201 ATGATACCGGAGACCGGCTCACCGCTCCAGATTTATCAGCAATAAACCAGCGACCC 7260  
7201 ATGATACCGGAGACCGGCTCACCGCTCCAGATTTATCAGCAATAAACCAGCGACCC 7260  
7261 GGAAGGGCGGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTTAAT 7320  
7261 GGAAGGGCGGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTTAAT 7320  
7321 TGTTCGGGAGCTAGAGTAAAGTATGCGGAGTAAATAGTTTGGCAGAGTTGTTGCC 7380  
7321 TGTTCGGGAGCTAGAGTAAAGTATGCGGAGTAAATAGTTTGGCAGAGTTGTTGCC 7380  
7381 ATTGTCGAGGACATCGTGGTGTGATGCTGCTGTTTGGTATGGTCTTATTCAGCTCGGT 7440

7381 ATTGTCGAGGACATCGTGGTGTGATCGCTCGTCTGTTGGTATGCTTCAATTCAGCTCGGT 7440  
7441 TCCCAAGCATCAAGGCGAGTTATCATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCC 7500  
7441 TCCCAAGCATCAAGGCGAGTTATCATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCC 7500  
7501 TTCCGCTCTCCGATCGTTGTGAGAAAGTTGGCGCGAGTGTATCACTCATGTTATG 7560  
7501 TTCCGCTCTCCGATCGTTGTGAGAAAGTTGGCGCGAGTGTATCACTCATGTTATG 7560  
7561 GCAGCACTGCATAATTTCTTACTGTATGTCATCCGATCCGTAAGATGCTTTCTGTGACTGT 7620  
7561 GCAGCACTGCATAATTTCTTACTGTATGTCATCCGATCCGTAAGATGCTTTCTGTGACTGT 7620  
7621 GAGTACTCTCAACCAAGTCAATCTGAGAAATAGTGTATGGCGAGCCAGTGTCTTCCCGC 7680  
7621 GAGTACTCTCAACCAAGTCAATCTGAGAAATAGTGTATGGCGAGCCAGTGTCTTCCCGC 7680  
7681 GCCTCAACACGGGATTAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCTGGA 7740  
7681 GCCTCAACACGGGATTAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCTGGA 7740  
7741 AAACGTTCTTCCGGGGGAAACCTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATG 7800  
7741 AAACGTTCTTCCGGGGGAAACCTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATG 7800  
7801 TAACCCACTCTGTCACCCAACTGATCTTCAAGATCTTACTTACCCAGCGTTTCTGGG 7860  
7801 TAACCCACTCTGTCACCCAACTGATCTTCAAGATCTTACTTACCCAGCGTTTCTGGG 7860  
7861 TGAGCAAAACACGAGGAGGAAATGCGCAAAAAGGGAATAGCGGACACGGAATCT 7920  
7861 TGAGCAAAACACGAGGAGGAAATGCGCAAAAAGGGAATAGCGGACACGGAATCT 7920  
7921 TGAATACTCATCTCTCTCTTTTCAATATTTAAGCAATTTATCAGGGTTATTTCTCTC 7980  
7921 TGAATACTCATCTCTCTCTTTTCAATATTTAAGCAATTTATCAGGGTTATTTCTCTC 7980  
7981 ATGAGCGGATACATATTTGAATGTTTGAATAAATAAACAATAGGGGTTCCGCCACA 8040  
7981 ATGAGCGGATACATATTTGAATGTTTGAATAAATAAACAATAGGGGTTCCGCCACA 8040  
8041 TTTCCCGGAAAGTGCACCTGACGCTTAAGAAACCATTTATCATGACATTAACCTAT 8100  
8041 TTTCCCGGAAAGTGCACCTGACGCTTAAGAAACCATTTATCATGACATTAACCTAT 8100  
8101 AAAAATAGCGTATACAGAGGCGCTTTCCTTCAAGAAATTAATTCATGTTTGCACG 8160  
8101 AAAAATAGCGTATACAGAGGCGCTTTCCTTCAAGAAATTAATTCATGTTTGCACG 8160  
8161 TTATCATGATAGCTGACTCATGTTGGTATTTGTAATAGAGCGAGATCGGGAACACTG 8220  
8161 TTATCATGATAGCTGACTCATGTTGGTATTTGTAATAGAGCGAGATCGGGAACACTG 8220  
8221 AAAAATAACAGTTATTATTTCG 8241  
8221 AAAAATAACAGTTATTATTTCG 8241

RESULT 2  
AAV28852  
ID AAV28852 standard; cDNA to mRNA; 8598 BP.  
XX  
AC AAV28852;  
XX  
DT 04-AUG-1998 (first entry)  
XX  
DE pPic9/ELF25L cDNA construct containing the elafin 25L gene.  
XX  
KW Elafin; elafin 25L gene; Pichia pastoris; expression vector;  
XX  
KW alcohol oxidase 1; controlling region; promoter; ds.  
XX  
OS Synthetic.



OS Pichia pastoris.  
 XX Key Location/Qualifiers  
 PH CDS /tag= a  
 FT 949..1377  
 FT sig\_peptide /tag= b  
 FT 949..1203  
 FT mat\_peptide /tag= c  
 FT 1204..1374  
 XX JP10127292-A.  
 PN 19-MAY-1998.  
 PD 31-OCT-1996; 96JP-00304233.  
 XX 31-OCT-1996; 96JP-00304233.  
 XX (TSUR ) TSUMURA & CO.  
 PA WPI; 1998-340657/30.  
 XX P-PSDB; AAM57237.  
 DR New elafin expression vector - used to transform microbial host(s) for  
 PT production of elafin commercially.  
 XX Example 1; Page 11-15; 21pp; Japanese.  
 XX The present sequence represents a pPIC9/ELF25L cDNA construct containing  
 CC the elafin 25L gene, used in an example of the present invention. The  
 CC present invention describes: (f) an elafin (EL) expression vector  
 CC comprising a gene encoding EL and a gene encoding alcohol oxidase 1  
 CC controlling region (AOCR) which controls the expression of EL encoding  
 CC gene; (2) a microbial host transformed with the vector; (3) DNA fragment  
 CC for recombinant transformation comprising: (a) a first region homologous  
 CC to the genomic DNA of the host; (b) a promoter region of (AOCR) gene; (c)  
 CC a gene encoding signal peptide; (d) a linker designed to secrete a  
 CC matured elafin; (e) a gene encoding EL; (f) selection marker gene; (g)  
 CC stop codon, and (h) a second region homologous to a part of the genomic  
 CC DNA of the host, and (4) a microbial host transformed with the DNA  
 CC fragment of (3). The microbial hosts can be used to prepare EL  
 CC commercially. The production of EL by the microbial hosts can be carried  
 CC on a large scale  
 XX Sequence 8598 BP; 2306 A; 2025 C; 1920 G; 2347 T; 0 U; 0 Other;  
 SQ Query Match 91.7%; Score 7555.6; DB 2; Length 8598;  
 Best Local Similarity 94.0%; Pred. No. 0;  
 Matches 8098; Conservative 0; Mismatches 129; Indels 385; Gaps 8;  
 QY 1 AGATCTACATCCAAAGACGAAAGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60  
 DB |||||  
 DB 1 AGATCTACATCCAAAGACGAAAGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60  
 QY 61 GTCCATTCTCACATAAGTGCCTCAACGCAACAGGAGGGATACACTAGCAGACACCGT 120  
 DB |||||  
 DB 61 GTCCATTCTCACATAAGTGCCTCAACGCAACAGGAGGGATACACTAGCAGACACCGT 120  
 QY 121 TCGCAACGAGACCTCCACCTCTCTCTCTCTCAACCCACCTTTTGGCATCCGACAAACC 180  
 DB |||||  
 DB 121 TCGCAACGAGACCTCCACCTCTCTCTCTCTCAACCCACCTTTTGGCATCCGACAAACC 180  
 QY 181 AGCCAGATTATTGGCTTTGATTGGAGCTCGCTCAATTCCTCTATTAGGCTACTA 240  
 DB |||||  
 DB 181 AGCCAGATTATTGGCTTTGATTGGAGCTCGCTCAATTCCTCTATTAGGCTACTA 240  
 QY 241 ACACCATGACTTATTAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
 DB |||||  
 DB 241 ACACCATGACTTATTAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
 QY 301 TTTCGGAATGCAACGACTCCGCTATTACACCGCAACATCACTCCAGATGAGGCTTTCTG 360  
 DB |||||  
 DB 301 TTTCGGAATGCAACGACTCCGCTATTACACCGCAACATCACTCCAGATGAGGCTTTCTG 360

QY 361 AGTGTGGGGTCAAAATAGTTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420  
 DB |||||  
 DB 361 AGTGTGGGGTCAAAATAGTTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420  
 QY 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAATAGTTTGGTTCCG 480  
 DB |||||  
 DB 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAATAGTTTGGTTCCG 480  
 QY 481 TTGAAATGCTTAACCGCGAGTTGGTCAAAAAGAACTTCCAAAGTGGCCATACCGTTTCT 540  
 DB |||||  
 DB 481 TTGAAATGCTTAACCGCGAGTTGGTCAAAAAGAACTTCCAAAGTGGCCATACCGTTTCT 540  
 QY 541 CTTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAAATGCTTAGCGAGTCT 600  
 DB |||||  
 DB 541 CTTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAAATGCTTAGCGAGTCT 600  
 QY 601 CTCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAAGCGAAATGGGAAACACCCCGCT 660  
 DB |||||  
 DB 601 CTCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAAGCGAAATGGGAAACACCCCGCT 660  
 QY 661 TTTTGGATGATTATGCAATGCTTCCACATTTGATGCTTCCAAAGATTTCTGGTGGGAATCT 720  
 DB |||||  
 DB 661 TTTTGGATGATTATGCAATGCTTCCACATTTGATGCTTCCAAAGATTTCTGGTGGGAATCT 720  
 QY 721 GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780  
 DB |||||  
 DB 721 GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780  
 QY 781 ATATAACAGAGGAAGCTGCTGCTTTAAACCTTTTTTTTATCATCATTTAGCTT 840  
 DB |||||  
 DB 781 ATATAACAGAGGAAGCTGCTGCTTTAAACCTTTTTTTTATCATCATTTAGCTT 840  
 QY 841 ACTTTCATTAATGGGACTGGTCCCAATGACAGCTTTTGAATTTAAACGCTTTTAAACGA 900  
 DB |||||  
 DB 841 ACTTTCATTAATGGGACTGGTCCCAATGACAGCTTTTGAATTTAAACGCTTTTAAACGA 900  
 QY 901 CAACTTGAGAGATCAAAAACCACTAATTAATTCGAAGGATCCAAACGATGAGATTTCT 960  
 DB |||||  
 DB 901 CAACTTGAGAGATCAAAAACCACTAATTAATTCGAAGGATCCAAACGATGAGATTTCT 960  
 QY 961 TCAATTTTACTGAGTTTATTCGCGATCTCTCGCATTTAGCTGCTCCAGTCAACACT 1020  
 DB |||||  
 DB 961 TCAATTTTACTGAGTTTATTCGCGATCTCTCGCATTTAGCTGCTCCAGTCAACACT 1020  
 QY 1021 ACAACGAGATGAACCGGACAAATTCGGGTGAAGCTGTCTATCGTTTACTCAGATTTA 1080  
 DB |||||  
 DB 1021 ACAACGAGATGAACCGGACAAATTCGGGTGAAGCTGTCTATCGTTTACTCAGATTTA 1080  
 QY 1081 GAAGGGATTTCCGATGTTGTTTCCCATTTTCCACAGCACAATAACGGTTATTG 1140  
 DB |||||  
 DB 1081 GAAGGGATTTCCGATGTTGTTTCCCATTTTCCACAGCACAATAACGGTTATTG 1140  
 QY 1141 TTTTAAATATCTACTATTGCCAGCATTTGCTGTCTTAAAGAAAGAGGGGTATCTCTCGAGAA 1200  
 DB |||||  
 DB 1141 TTTTAAATATCTACTATTGCCAGCATTTGCTGTCTTAAAGAAAGAGGGGTATCTCTCGAGAA 1200  
 QY 1201 AGAG---AGGCTGAAGCCCGGAGCCCGGAGCCCGGAGCCCGGAGGAAACATGATC 1257  
 DB |||||  
 DB 1201 AGAGCTCAAGAACCAAGTTAAGGGTCCGGTGTGCAACAAACCGGGCTCTTGGCCGATTATC 1260  
 QY 1258 ATGCCCCCGGGAGAGGATPAAATTTGGGTTCCTCGGTGTCCCGGCTGTCACCGCCAGAGTCCACG 1317  
 DB |||||  
 DB 1261 CTGATCGGCTGGCTTTGCTGAACCCCGGAAACCGTTGTCTGAAAGACACTGACTGACCCG 1320  
 QY GAG----- 1318  
 DB GGTATCAAAAAATGCTGCGAAGGTTCTTCGGGTATGGCATGCTTGTTCGCGAGTAGTGA 1380  
 QY ----- 1321  
 DB AGCTTCGAGGCCAGCTTGGCTGAGGTGCGAAGAGAACCCATCTGCCTAATGAGTAA 1440

QY 1321 ----- 1320  
Db 1441 TATCTGATAAGAACGCTTTTATTTAGCTTTTAACTACATAGACGAATAATAATACGA 1500  
QY 1321 ----- 1320  
Db 1501 TCGGGATACATTAGCATATTGAGTCTTTTACCTCTCAATTTGCTCAAGATTAGTATCGT 1560  
QY 1321 ----- 1320  
Db 1561 GACCAATTTTAGATCATTTTCGGTCAGACACGCCACATTTTCGCGTAATTTTAAAGTTCAA 1620  
QY 1321 -----AGAGTTGCTGTTTTGATGACAGTCTCGGGGAT 1354  
Db 1621 TTAATTGATGCGTGACACACTATATTAGCGTCAACTGTTTAATAAGGAAGTAGTGATAAT 1680  
QY 1355 TCCCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGA-----AGAAGAAATGTC 1409  
Db 1681 AACGGTTGAGTTGCTTCAATGCTTATTTCAAAATCTAAGATGTTTAAACATTTTGG 1740  
QY 1410 CTTCTAACTAGTGGCGT-----AGAA 1430  
Db 1741 ATACTAACGACGATAGTTCTCCTCGCATCTGCCACCGTTGATATTAGTAAACTACAAGAA 1800  
QY 1431 TTCCTTAGGGCGCGCAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTG 1490  
Db 1801 TTCCTTAGGGCGCGCGAAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTG 1860  
QY 1491 GGCACCTTAGGAGAACCGGCTTCTGCTAGATTCTAATCAAGAGGATGTCAGAAATGCCATT 1550  
Db 1861 GGCACCTTAGGAGAACCGGCTTCTGCTAGATTCTAATCAAGAGGATGTCAGAAATGCCATT 1920  
QY 1551 TGCCTGAGAGATGAGGCTTCATTTTGATGACTTTTATTTGTAACCTATATAGTATAG 1610  
Db 1921 TGCCTGAGAGATGAGGCTTCATTTTGATGACTTTTATTTGTAACCTATATAGTATAG 1980  
QY 1611 GATTTTTTTTGCTATTTGTTCTCTCTGACGAGCTTCTCCTGATCAGGCTATCTCGC 1670  
Db 1981 GATTTTTTTTGCTATTTGTTCTCTCTGACGAGCTTCTCCTGATCAGGCTATCTCGC 2040  
QY 1671 AGCTGATGAATATCTTTGTTAGGGGTTTGGAAATCATTTGAGTTGATGTTTTCTT 1730  
Db 2041 AGCTGATGAATATCTTTGTTAGGGGTTTGGAAATCATTTGAGTTGATGTTTTCTT 2100  
QY 1731 GGTATTTCCACTCTCTTTCAGAGTACAGAGATTAACTGAGAGAGTTGTTGTTGCAAGC 1790  
Db 2101 GGTATTTCCACTCTCTTTCAGAGTACAGAGATTAACTGAGAGAGTTGTTGTTGCAAGC 2160  
QY 1791 TTATCGATAAGCTTTAATGCGTAGTTTATCACAGTTAAATGCTAACGCACTCAGGCAC 1850  
Db 2161 TTATCGATAAGCTTTAATGCGTAGTTTATCACAGTTAAATGCTAACGCACTCAGGCAC 2220  
QY 1851 CGTGTATGAATCTAAGATGCGCTCATCTCGGACCGTCACTCCCTGATGCTG 1910  
Db 2221 CGTGTATGAATCTAAGATGCGCTCATCTCGGACCGTCACTCCCTGATGCTG 2280  
QY 1911 TAGGCATAGGCTTGGTTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGG 1970  
Db 2281 TAGGCATAGGCTTGGTTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGG 2340  
QY 1971 ACAGCATCGCCAGTCACTATGCGGTGCTGTAGCGCTATATGCGTTGATGCAATTTCTAT 2030  
Db 2341 ACAGCATCGCCAGTCACTATGCGGTGCTGTAGCGCTATATGCGTTGATGCAATTTCTAT 2400  
QY 2031 GGGCACCGCTTCTCGGACGACTGTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTT 2090  
Db 2401 GGGCACCGCTTCTCGGACGACTGTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTT 2460  
QY 2091 CGCTACTTTGGAGCCACTATPCGACTAGCGATCATGGGACACACCCGCTCTGTGGATCT 2150  
Db 2461 CGCTACTTTGGAGCCACTATPCGACTAGCGATCATGGGACACACCCGCTCTGTGGATCT 2520  
QY 2151 ATCGAATCTAATGTAAGTTAAATCTCTAAATTAATTAAGTCCCGCTTTCTCCATA 2210

Db 2521 ATCGAATCTAATGTAAGTTAAATCTCTAAATTAATTAAGTCCCGCTTTCTCCATA 2580  
QY 2211 CGAACCTTAAACAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACT 2270  
Db 2581 CGAACCTTAAACAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACT 2640  
QY 2271 GCTTGCCCAATATGTTTTCAGTCCCTCAGGAGTTACGTTCTGTGAAAGTGAATCTCTGG 2330  
Db 2641 GCTTGCCCAATATGTTTTCAGTCCCTCAGGAGTTACGTTCTGTGAAAGTGAATCTCTGG 2700  
QY 2331 AAGGTTGACGAGTTAACTCCGCTGATTCACGGGCAATATCCGTACGTTGCGAAAGTCTGG 2390  
Db 2701 AAGGTTGACGAGTTAACTCCGCTGATTCACGGGCAATATCCGTACGTTGCGAAAGTCTGG 2760  
QY 2391 TTGCTACCCGAGGAGTAATCTCCAACTCTCTGGAGAGTAGGCACCAACAACACAGAT 2450  
Db 2761 TTGCTACCCGAGGAGTAATCTCCAACTCTCTGGAGAGTAGGCACCAACAACACAGAT 2820  
QY 2451 CAGCGTGTGTTACTTGTATCAATAGAAAGCAATTCGATTTGCAAGGATCAAGTCT 2510  
Db 2821 CAGCGTGTGTTACTTGTATCAATAGAAAGCAATTCGATTTGCAAGGATCAAGTCT 2880  
QY 2511 TCAGGAGCGTACTGATTTGGACATTTCCAAAGCCTGCTCTAGGTTGCAACCGATAGGGTT 2570  
Db 2881 TCAGGAGCGTACTGATTTGGACATTTCCAAAGCCTGCTCTAGGTTGCAACCGATAGGGTT 2940  
QY 2571 GTAGAGTGTCAATACATCTGCGTCAATTTCAAACCTTTGGCAACCTGCAAGCTTGGTTG 2630  
Db 2941 GTAGAGTGTCAATACATCTGCGTCAATTTCAAACCTTTGGCAACCTGCAAGCTTGGTTG 3000  
QY 2631 TGAACAGCATCTTCAATTTCTGGCAAGCTCTCTGCTCATATCGACAGCACAACAATC 2690  
Db 3001 TGAACAGCATCTTCAATTTCTGGCAAGCTCTCTGCTCATATCGACAGCACAACAATC 3060  
QY 2691 ACCTGGGAATCAATACCATTTGAGCTTGAGACAGAAAGTCTGAGCAACGAAATCTGGA 2750  
Db 3061 ACCTGGGAATCAATACCATTTGAGCTTGAGACAGAAAGTCTGAGCAACGAAATCTGGA 3119  
QY 2751 TCAGCGTATTTATCAGCAATTAAGTCACTTCAGAGGGCCAGCAGGCAATGTCANACTA 2810  
Db 3120 TCAGCGTATTTATCAGCAATTAAGTCACTTCAGAGGGCCAGCAGGCAATGTCANACTA 3179  
QY 2811 CACAGGCGTATGTTGCTCATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGTTTCT 2870  
Db 3180 CACAGGCGTATGTTGCTCATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGTTTCT 3239  
QY 2871 GGAACCAATATTTTGTCACTTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGTACT 2930  
Db 3240 GGAACCAATATTTTGTCACTTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGTACT 3299  
QY 2931 GCCTGGCGGCTCTCTCTAGCAGATACACTTAGCAACCACTTGTGGCAACGTAAGTG 2990  
Db 3300 GCCTGGCGGCTCTCTCTAGCAGATACACTTAGCAACCACTTGTGGCAACGTAAGTG 3359  
QY 2991 ACTTCTGGGTAAAGGTACCATCTTCTTAGGTGGAGATGCAAAAAACAATTTCTTTGCAA 3050  
Db 3360 ACTTCTGGGTAAAGGTACCATCTTCTTAGGTGGAGATGCAAAAAACAATTTCTTTGCAA 3419  
QY 3051 CAGCAACTTTGCGAGCAACACCCAGCATCAGGAGTGGAGGCAAGATTCGCGTTCCA 3110  
Db 3420 CAGCAACTTTGCGAGCAACACCCAGCATCAGGAGTGGAGGCAAGATTCGCGTTCCA 3479  
QY 3111 CCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTTCGAAAAACGAGAGCAGACTACACCA 3170  
Db 3480 CCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTTCGAAAAACGAGAGCAGACTACACCA 3539  
QY 3171 GGGCAAGTCTCAACTTTGCAACGCTCTCGTTAGTCTCATGGAAATTTCTGAGGTTA 3230  
Db 3540 GGGCAAGTCTCAACTTTGCAACGCTCTCGTTAGTCTCATGGAAATTTCTGAGGTTA 3599  
QY 3231 TCTATAGAGAGTCAATGGCTCTCTTAAAGCTTATCTGGCAATTCGCAATAGTTCTCTCTGG 3290

Db	3600	TCATAGAGAGATCAATGGGCTCTCTTAAAGTTATCTGGCAATTCGATAAGTTCTCTGGG	3659	QY	4370	GATACTAGCATTAGCGGCAAGGATGCAACTTTATCAACGAGGTCCTATAGATAACCCCT	4429
QY	3291	AAAGGAGCTTCTAACACAGGTGCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCT	3350	Db	4739	GATACTAGCATTAGCGGCAAGGATGCAACTTTATCAACGAGGTCCTATAGATAACCCCT	4798
Db	3660	AAAGGAGCTTCTAACACAGGTGCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCT	3719	QY	4430	AGCGCTGGGATCATCTTTGGACAACTCTTTTCGCCAAATCTAGTCCAAATCACTTC	4489
QY	3351	AAAAGGGCTTTGTCCACATTTGACGAACTTTGCGAATTCGTTGACTAATCCATA	3410	Db	4799	AGCGCTGGGATCATCTTTGGACAACTCTTTTCGCCAAATCTAGTCCAAATCACTTC	4858
Db	3720	AAAAGGGCTTTGTCCACATTTGACGAACTTTGCGAATTCGTTGACTAATCCATA	3779	QY	4490	ATTGATACCATTTATGTTACAACTTTGAGCAAGTTGTCGATCAGCTCTCAAAATTGGTCTCTC	4549
QY	3411	ATCTGTTCCGTTTCTGGATAGGACGAGGAGGCTCTCAATTTCTTGTGAGGAGGCC	3470	Db	4859	ATTGATACCATTTATGTTACAACTTTGAGCAAGTTGTCGATCAGCTCTCAAAATTGGTCTCTC	4918
Db	3780	ATCTGTTCCGTTTCTGGATAGGACGAGGAGGCTCTCAATTTCTTGTGAGGAGGCC	3839	QY	4550	TGTAACGAGTACCACTTGCACATTAATTTGAAGCTCAGTCGATTTGAGTGAACCTTGAT	4609
QY	3471	TTAGAAACGCTCAATTTTGCAAAATTCATACGACCTTTCAGAGGAGCTCTTTAGGTTTG	3530	Db	4919	TGTAACGAGTACCACTTGCACATTAATTTGAAGCTCAGTCGATTTGAGTGAACCTTGAT	4978
Db	3840	TTAGAAACGCTCAATTTTGCAAAATTCATACGACCTTTCAGAGGAGCTCTTTAGGTTTG	3899	QY	4610	CAGGTTGTGCACTGGTCAGCAGCATAGGGAACAACGCGCTTTTCTACCAAACTCAAGGA	4669
QY	3531	GATTCTCTTTAGGTTGTTCTGGTGATCTCCGCTTGGCATCTCTTTCTCTCTAGTG	3590	Db	4979	CAGGTTGTGCACTGGTCAGCAGCATAGGGAACAACGCGCTTTTCTACCAAACTCAAGGA	5038
Db	3900	GATTCTCTTTAGGTTGTTCTGGTGATCTCCGCTTGGCATCTCTTTCTCTCTAGTG	3959	QY	4670	ATTATCAAACTCTGCAACACTTGGTATGAGGTAGCAAGGGAATGTCATCTTGAAGT	4729
QY	3591	ACCTTTAGGAGCTTCATATCCAGGTTTCTCTCCACCTCGTCCAACTGTCACCGTACTTG	3650	Db	5039	ATTATCAAACTCTGCAACACTTGGTATGAGGTAGCAAGGGAATGTCATCTTGAAGT	5098
Db	3960	ACCTTTAGGAGCTTCATATCCAGGTTTCTCTCCACCTCGTCCAACTGTCACCGTACTTG	4019	QY	4730	CGGACAGTGAGTTAGTCTTGAAGAAATTTCTGAAGCCGTATTTTATTATCAGTGAGTCAG	4789
QY	3651	GCACTCTAACTAATGCAAAATAAATAAGTACGACATTTCCAGGCTATATCTTCTCTTG	3710	Db	5099	CGGACAGTGAGTTAGTCTTGAAGAAATTTCTGAAGCCGTATTTTATTATCAGTGAGTCAG	5158
Db	4020	GCACTCTAACTAATGCAAAATAAATAAGTACGACATTTCCAGGCTATATCTTCTCTTG	4079	QY	4790	TCATCAGAGATCTCTACGCGGACGATCGTGGCGGACCTCGAGGTCGGCATCACGG	4849
QY	3711	GATTAGTCTCGAAGTTGATCAGCTTCTCCCTCAATTTTAGGCTTC-AACAAACTTC	3769	Db	5159	TCATCAGAGATCTCTACGCGGACGATCGTGGCGGACCTCGAGGTCGGCATCACGG	5206
Db	4080	GATTAGTCTCGAAGTTGATCAGCTTCTCCCTCAATTTTAGGCTTC-AACAAACTTC	4139	QY	4850	CGCACAGTGGGTTGCTGGCGCTATATCGCGCATACCCGATGGGGAAGATCGGGC	4909
QY	3770	GTCGCTCAAAATACCGTTTGGTATAGAAACCTTTCGGAGCATTCCTTTCAGTCCCAACAA	3829	Db	5207	CGCACAGTGGGTTGCTGGCGCTATATCGCGCATACCCGATGGGGAAGATCGGGC	5266
Db	4140	GTCGCTCAAAATACCGTTTGGTATAGAAACCTTTCGGAGCATTCCTTTCAGTCCCAACAA	4199	QY	4910	TCGCGCACTTCGCGCTCATAGCGCTTGTTCGCGGTGGGTATGTTGGCAGCCCGCTGGC	4969
QY	3830	GGTGGCTTCCATGGCTTAAGACCTTTGATGTCGCAAAACAGAGTGGTTCAGAGTG	3889	Db	5267	TCGCGCACTTCGCGCTCATAGCGCTTGTTCGCGGTGGGTATGTTGGCAGCCCGCTGGC	5326
Db	4200	GGT-GCTTCCATGGCTTAAGACCTTTGATGTCGCAAAACAGAGTGGTTCAGAGTG	4258	QY	4970	CGGCGGACTTTGGCGGCACTCTCTTCGATGCACTTCCTCGCGCGGCGGTCTCAA	5029
QY	3890	ACAGAAACCAACCTGTTTCTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAA	3949	Db	5327	CGGCGGACTTTGGCGGCACTCTCTTCGATGCACTTCCTCGCGCGGCGGTCTCAA	5386
Db	4259	ACAGAAACCAACCTGTTTCTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAA	4318	QY	5030	CGGCGCTCAACCTACTACTCTGGGCTGCTTCTTAATGCAAGAGTCGATTAAGGAGAGCGTCG	5089
QY	3950	TTGATACCCAGCAACTTTTGGTGGTCTGAGATGTCAGATGTCAGACCTTTATACCAAAACCGTG	4009	Db	5387	CGGCGCTCAACCTACTACTCTGGGCTGCTTCTTAATGCAAGAGTCGATTAAGGAGAGCGTCG	5446
Db	4319	TTGATACCCAGCAACTTTTGGTGGTCTGAGATGTCAGATGTCAGACCTTTATACCAAAACCGTG	4378	QY	5090	AGTATCTATGATGGAAGTATGGGATGTTGATACCGCATCTCTTCAGTGCTTTGAGGTC	5149
QY	4010	ACGAGGATTTGGTAGACTCCAGTTTGTCTCTTATAGCCTCCGGAATAGACTTTTGGGA	4069	Db	5447	AGTATCTATGATGGAAGTATGGGATGTTGATACCGCATCTCTTCAGTGCTTTGAGGTC	5506
Db	4379	ACGAGGATTTGGTAGACTCCAGTTTGTCTCTTATAGCCTCCGGAATAGACTTTTGGGA	4438	QY	5150	TCCTATCAGATTTATCCCAACTTAAAGCAACCGGAGGAGAGATTTTCATGTAATTTCTC	5209
QY	4070	CGAGTACACCGGCGCAACGAGTAAATTAGAGAGTCCAGCCCAAAAGTATGATAGACC	4129	Db	5507	TCCTATCAGATTTATCCCAACTTAAAGCAACCGGAGGAGAGATTTTCATGTAATTTCTC	5566
Db	4439	CGAGTACACCGGCGCAACGAGTAAATTAGAGAGTCCAGCCCAAAAGTATGATAGACC	4498	QY	5210	TCGATTTTGGTCACTCAGTAGACTGGAAGTGGAGTATCTCGGTTATGACAGAGAAAT	5269
QY	4130	ATCGGGGCGGTGAGTCAAGAGCGCAACAAATTTCTAGCAGGGAACCTTTTGGAC	4189	Db	5567	TCGATTTTGGTCACTCAGTAGACTGGAAGTGGAGTATCTCGGTTATGACAGAGAAAT	5626
Db	4499	ATCGGGGCGGTGAGTCAAGAGCGCAACAAATTTCTAGCAGGGAACCTTTTGGAC	4558	QY	5270	GTCTCTTCTGGAGACAGTAAATGAAAGTCCCACTAAAGAAATTCCTTTGTTATCAGGAAC	5329
QY	4190	ATCTTTAGAAAGTTCGTATTCAGTAGTCAATTTGCGAGTCAATTAATGGGATATATACC	4249	Db	5627	GTCTCTTCTGGAGACAGTAAATGAAAGTCCCACTAAAGAAATTCCTTTGTTATCAGGAAC	5686
Db	4559	ATCTTTAGAAAGTTCGTATTCAGTAGTCAATTTGCGAGTCAATTAATGGGATATATACC	4618	QY	5330	AAACTTTCTTTTTCGAACTTTTTCGGTGCCTTTGAACTATAAAATGAGTGGATATGTC	5389
QY	4250	AGAAGCAACGTGGAGTCACTATACCACTTTTGGGCTTCAGAAAGCATAAACAGT	4309	Db	5687	AAACTTTCTTTTTCGAACTTTTTCGGTGCCTTTGAACTATAAAATGAGTGGATATGTC	5746
Db	4619	AGAAGCAACGTGGAGTCACTATACCACTTTTGGGCTTCAGAAAGCATAAACAGT	4678	QY	5390	GGGTAGGATGAGCGGCAATGCTTACCTTCTGACCTTCAAGAGGTATGAGGGTTT	5449
QY	4310	TCTACTACCGCATTTAGTGAACCTTTTCAAAATCGCCCAAGTGGAGAAAGGACACAGC	4369	Db	5747	GGGTAGGATGAGCGGCAATGCTTACCTTCTGACCTTCAAGAGGTATGAGGGTTT	5806
Db	4679	TCTACTACCGCATTTAGTGAACCTTTTCAAAATCGCCCAAGTGGAGAAAGGACACAGC	4738				

[illegible]



QY 61 GTCCATTCTCACATAAGTGCACAAACGACAGAGGGGATACACTAGCAGCAGACCGT 120  
 Db 61 GTCCATTCTCACATAAGTGCACAAACGACAGAGGGGATACACTAGCAGCAGACCGT 120  
 QY 121 TGCACACGAGGACCTCCACTCTCTCTCTCTCAACACCCACTTTTCCATCGAATAACC 180  
 Db 121 TGCACACGAGGACCTCCACTCTCTCTCTCTCAACACCCACTTTTCCATCGAATAACC 180  
 QY 181 AGCCGAGTTATGGGCTTGAATGGAGCTCGCTCATTTCCAAATTCCTTATTAGGCTACTA 240  
 Db 181 AGCCGAGTTATGGGCTTGAATGGAGCTCGCTCATTTCCAAATTCCTTATTAGGCTACTA 240  
 QY 241 ACACATGACTTTATTAGCTGTCTATCTGCGCCCTCTGCGAGGTTTCATGTTGTTTA 300  
 Db 241 ACACATGACTTTATTAGCTGTCTATCTGCGCCCTCTGCGAGGTTTCATGTTGTTTA 300  
 QY 301 TTTCCGAATGCAACAAAGCTCGCATTCACACCGCAACATCACTCCAGATGAGGGCTTTCTG 360  
 Db 301 TTTCCGAATGCAACAAAGCTCGCATTCACACCGCAACATCACTCCAGATGAGGGCTTTCTG 360  
 QY 361 AGTGTGGGTCAAATAGTTTCATGTTCCCAAAATGGCCCAAAATGACAGATTTAAACGCT 420  
 Db 361 AGTGTGGGTCAAATAGTTTCATGTTCCCAAAATGGCCCAAAATGACAGATTTAAACGCT 420  
 QY 421 GTCTTGGAACTTAATGACAAAAGCGTGATCTCATCCAAAGATGAATCAAGTTTGGTTTCG 480  
 Db 421 GTCTTGGAACTTAATGACAAAAGCGTGATCTCATCCAAAGATGAATCAAGTTTGGTTTCG 480  
 QY 481 TTGAATGCTTAACGCCCTAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCATACCGTTTGT 540  
 Db 481 TTGAATGCTTAACGCCCTAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCATACCGTTTGT 540  
 QY 541 CTGTTTGGTATTGATGACGAATGCTCAAAAATATCTCATTAATGCTTAGCGAGTCT 600  
 Db 541 CTGTTTGGTATTGATGACGAATGCTCAAAAATATCTCATTAATGCTTAGCGAGTCT 600  
 QY 601 CTCTATCGCTTCTGAACCCCGTGACCTGTGCCGAAACGCAAAATGGGGAACACCCGCT 660  
 Db 601 CTCTATCGCTTCTGAACCCCGTGACCTGTGCCGAAACGCAAAATGGGGAACACCCGCT 660  
 QY 661 TTTTGGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 Db 661 TTTTGGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 QY 721 GCTGATAGCTAAAGTTCATGATCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 780  
 Db 721 GCTGATAGCTAAAGTTCATGATCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 780  
 QY 781 ATATAACAGAGGAGCTGCCCTGCTCTTAAACCTTTTTTTTATCATCATTAATAGCTT 840  
 Db 781 ATATAACAGAGGAGCTGCCCTGCTCTTAAACCTTTTTTTTATCATCATTAATAGCTT 840  
 QY 841 ACTTTTCATTAATGCGACTGGTTCCAAATGACAGCTTTTGAATTAACGACTTTTAACGA 900  
 Db 841 ACTTTTCATTAATGCGACTGGTTCCAAATGACAGCTTTTGAATTAACGACTTTTAACGA 900  
 QY 901 CAACTTGAGAAGATCAAAAACAACTAAATTTTGAAGGATCCAAACGATGAGATTTTCT 960  
 Db 901 CAACTTGAGAAGATCAAAAACAACTAAATTTTGAAGGATCCAAACGATGAGATTTTCT 960  
 QY 961 TCAATTTTACTGAGTTTATTCGAGATCCTCCGATTAAGTCTCCAGTCAAACACT 1020  
 Db 961 TCAATTTTACTGAGTTTATTCGAGATCCTCCGATTAAGTCTCCAGTCAAACACT 1020  
 QY 1021 ACAACAGAGATGAACGSCAAATTCGCGTGAAGCTGTCATCGGTTACTCAGATTTA 1080  
 Db 1021 ACAACAGAGATGAACGSCAAATTCGCGTGAAGCTGTCATCGGTTACTCAGATTTA 1080  
 QY 1081 GAAGGGATTTCAATGTTGCTGTTTGGCAATTTCCAAACAGCACAATAACCGGTTATTG 1140  
 Db 1081 GAAGGGATTTCAATGTTGCTGTTTGGCAATTTTCCAAACAGCACAATAACCGGTTATTG 1140

QY 1141 TTTTATAATACTACTATTGCCAGCATTTGCTGTCTTAAAGAAAGAGGGGTATCTCTCGAGAAA 1200  
 Db 1141 TTTTATAATACTACTATTGCCAGCATTTGCTGTCTTAAAGAAAGAGGGGTATCTCTCGAGAAA 1200  
 QY 1201 AGAGAGCTGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1236  
 Db 1201 AGAGAGCTGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1236  
 QY 1237 ----- 1236  
 Db 1237 ----- 1236  
 QY 1261 GTTGAGACTTCCCGCTCTGTATTTCACAGAACCGCCAAAGTTACTGAAGTTACAAACGCT 1320  
 Db 1237 ----- 1236  
 QY 1237 ----- 1236  
 Db 1321 ACCGGCATCGTTGATTGTTGGTTTCGAAAATCGCGCTTCAAAGGCCAAGAGACCTCGGTAAC 1380  
 QY 1237 ----- 1236  
 Db 1381 GGCTGAAAGCCATTTGGCAGGTTGAGCAAAAGCATCTATCGCCGCTACTGACTCCGCT 1440  
 QY 1237 ----- 1236  
 Db 1441 TGGGGCAACCGCCAACTCTTCATCGGCTTGAAGAGCGGCTTCGGTAAATTTGCGCGTCGCT 1500  
 QY 1237 -----GCCCAGGAAGAAACATGTATCATGGCCCGCCCGGAGAGGATAA----- 1279  
 Db 1501 CGTTTGAACAGCTCTCTGAAAGACACCGGCGACATCAATCCTTGGATAGCAAAAGCGAC 1560  
 QY 1280 ----- 1279  
 Db 1561 TATTTGGGTGTAACAAATTTGCCGAACCGCAGGACGCGCTCATTTCCGTAAGCTACGAT 1620  
 QY 1280 ----- 1279  
 Db 1621 TCTCCGAATTTCCGGCTCAGCGCAGGTTACAATAGCGGCTTAACGACATGCGAGC 1680  
 QY 1280 ----- 1279  
 Db 1681 AGACATAACAGCAATCTTACCACGCGGCTTCAACTACAAAAACGTTGGCTTCTTCGTG 1740  
 QY 1280 ----- 1279  
 Db 1741 CAATATGGCGGTGCTTATAAAGACATCATCAAGTGAAGAGGCTTGAATATTGAGAAA 1800  
 QY 1280 ----- 1279  
 Db 1801 TACCAAGTTACCGTTTGGTCAGCGGTTAGACAAATGATGCCCTGTACGCTTCCGTAGCC 1860  
 QY 1280 ----- 1279  
 Db 1861 GTACAGCAACAGACGCAAACTGACTGATGCTTCCAATTCGCACAACTCTCAAACCGAA 1920  
 QY 1280 -----ATTGTGCTTCCCGGTGTCAACCGCCAGC 1309  
 Db 1921 GTTCCGCTACCTTGGCATACCGCTTCGGCAAGCTAACCGCCGAGTTTCTTACGCGCAC 1980  
 QY 1310 AGTGCAACGAGAGGTTGCTGTTTGTATGACAGTGTCCG----- 1349  
 Db 1981 GGCTTCAAAGGTTTGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
 QY 1350 -----GGGATTCGCTGCTTCCACCCCATGCGCATCGA 1385  
 Db 2041 GTTCCGCAATACGACTTCTCCAAACGCACTTCTGCTTGGTTTCTGCCGTTGGTTGCAA 2100  
 QY 1386 GAACACTCAAGAAAGAAAGATGTCCTTCTAACTAGTGGG-----TA 1427  
 Db 2101 GAAGCAAGGCGCAAAATTCGTAGCGACTTCGCGCGGTTGTTGTTGCTGCGTCAAAA 2160  
 QY 1428 GAATTCCTTAGGCGCGCGGAAATTAATTCGCTTAGACATGCTGTTCTCTAGTTCAAG 1487  
 Db 2161 TTCTAACCTAGGCGCGCGGAAATTAATTCGCTTAGACATGCTGTTCTCTAGTTCAAG 2220  
 QY 1488 TTGGGCATTTACGAGAAGACCGGCTTCTGCTAGATTCTAATCAAGAGGATGTCAGAAATGCC 1547





QY 3708 TTGGATTAGCTTTCGCAAGTTTCATCAGCTTCTCCCTAAATTTTAGCGTTCAACAAACT 3767  
DB |||||  
QY 4441 TTGGATTAGCTTTCGCAAGTTTCATCAGCTTCTCCCTAAATTTTAGCGTTCAACAAACT 4500  
DB |||||  
QY 3768 TCGTCGTCAAAATACCGTTTGGTATAGAACTTCTGGAGCATTCCTTACGATCCAC 3827  
DB |||||  
QY 4501 TCGTCGTCAAAATACCGTTTGGTATAGAACTTCTGGAGCATTCCTTACGATCCAC 4560  
DB |||||  
QY 3828 AAGTGGCTTCCATGGCTTAAAGACCTTTGATTGGCCAAAACAGGAGTGCCTTCCAAG 3887  
DB |||||  
QY 4561 AAGTGGCTTCCATGGCTTAAAGACCTTTGATTGGCCAAAACAGGAGTGCCTTCCAAG 4620  
DB |||||  
QY 3888 TGACAGAAACCAACACCTTTGTTTCAACCAAAATTTCAAGCATCTCCATCAATCC 3947  
DB |||||  
QY 4621 TGACAGAAACCAACACCTTTGTTTCAACCAAAATTTCAAGCATCTCCATCAATCC 4680  
DB |||||  
QY 3948 AATTCGATACCCAGCAACTTTTGTAGTCTCCAGATGTAGCACTTTATATACCAAAACCG 4007  
DB |||||  
QY 4681 AATTCGATACCCAGCAACTTTTGTAGTCTCCAGATGTAGCACTTTATATACCAAAACCG 4740  
DB |||||  
QY 4008 TGACAGCAGATTTGTAGACTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGT 4067  
DB |||||  
QY 4741 TGACAGCAGATTTGTAGACTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGT 4800  
DB |||||  
QY 4068 GACGAGTACACAGCCCAACAGTAATAGAGAGTCAGCCACCAAGTAGTAGATA 4127  
DB |||||  
QY 4801 GACGAGTACACAGCCCAACAGTAATAGAGAGTCAGCCACCAAGTAGTAGATA 4860  
DB |||||  
QY 4128 CCATCGGGGGCTCAGTAGTCAAGACGCCCAAAATTTCACTGACAGGCACTTTTGT 4187  
DB |||||  
QY 4861 CCATCGGGGGCTCAGTAGTCAAGACGCCCAAAATTTCACTGACAGGCACTTTTGT 4920  
DB |||||  
QY 4188 ACATCTTCAGAAAGTTCGTATTCAGTAGTCAATTTGCCAGCATCAATATGGGATATA 4247  
DB |||||  
QY 4921 ACATCTTCAGAAAGTTCGTATTCAGTAGTCAATTTGCCAGCATCAATATGGGATATA 4980  
DB |||||  
QY 4248 CCAGAGCAACAGTGGAGTACACTTACCACTTTGGGTCTCAGAAAAGCATAACA 4307  
DB |||||  
QY 4981 CCAGAGCAACAGTGGAGTACACTTACCACTTTGGGTCTCAGAAAAGCATAACA 5040  
DB |||||  
QY 4308 GTTCTACTACGCCATTAGTGAACCTTTTCAATGCCCCAGTGGAGAGAAAAGGCACA 4367  
DB |||||  
QY 5041 GTTCTACTACGCCATTAGTGAACCTTTTCAATGCCCCAGTGGAGAGAAAAGGCACA 5100  
DB |||||  
QY 4368 GCGATCTAGCATTAGCGGCAAGATGCACTTTTCAACAGGGTCTTATAGTAACC 4427  
DB |||||  
QY 5101 GCGATCTAGCATTAGCGGCAAGATGCACTTTTCAACAGGGTCTTATAGTAACC 5160  
DB |||||  
QY 4428 CTAGCGCTGGGATCATCTTTGGACAACCTTTTCTGCAAACTTAGGTCCAAAATCACT 4487  
DB |||||  
QY 5161 CTAGCGCTGGGATCATCTTTGGACAACCTTTTCTGCAAACTTAGGTCCAAAATCACT 5220  
DB |||||  
QY 4488 TCATTGTACCAATTTAGTACAACTTGAGCAAGTTTGCATCAGCTCCAAATGGTCC 4547  
DB |||||  
QY 5221 TCATTGTACCAATTTAGTACAACTTGAGCAAGTTTGCATCAGCTCCAAATGGTCC 5280  
DB |||||  
QY 4548 TCTGTAAACGATGACTCAACTTGACATTAATTCAGCTCAGTCGATGAGTGAACCTTG 4607  
DB |||||  
QY 5281 TCTGTAAACGATGACTCAACTTGACATTAATTCAGCTCAGTCGATGAGTGAACCTTG 5340  
DB |||||  
QY 4608 ATCAGGTTGTGAGTGTGACAGCATAGGAAACACGGCTTTTCTTACCAAACTCAAG 4667  
DB |||||  
QY 5341 ATCAGGTTGTGAGTGTGACAGCATAGGAAACACGGCTTTTCTTACCAAACTCAAG 5400  
DB |||||  
QY 4668 GAATTTATCAACTCTGCAACCTTCGGTATCCAGTACAGGGAATGTCTACTTTGAA 4727  
DB |||||  
QY 5401 GAATTTATCAACTCTGCAACCTTCGGTATCCAGTACAGGGAATGTCTACTTTGAA 5460  
DB |||||  
QY 4728 GTCGACAGTGTAGTCTTGAGAAATTTCTGAAGCCGTATTTTATATCAGTGAATC 4787  
DB |||||  
QY 5461 GTCGACAGTGTAGTCTTGAGAAATTTCTGAAGCCGTATTTTATATCAGTGAATC 5520  
DB |||||

QY 4788 AGTCATCAGAGATCCTCTACGCCGACGCATCGTGGCCGACCTGCAGGTTCGCATCAC 4847  
DB |||||  
QY 5521 AGTCATCAGAGATCCTCTACGCCGACGCATCGTGGCCGACCTGCAGGTTCGCATCAC 5580  
DB |||||  
QY 4848 GCGCCACAGTGCCTTGTCTGGCCCTATATCGCCGACATCACCGATGGGAGATCG 4907  
DB |||||  
QY 5581 GCGCCACAGTGCCTTGTCTGGCCCTATATCGCCGACATCACCGATGGGAGATCG 5640  
DB |||||  
QY 4908 GCTCGCCACTTTCGGGCTCATGAGCGCTTGTTCGGCGTGGGTATGTCAGCCGCCGTG 4967  
DB |||||  
QY 5641 GCTCGCCACTTTCGGGCTCATGAGCGCTTGTTCGGCGTGGGTATGTCAGCCGCCGTG 5700  
DB |||||  
QY 4968 GCGGGGAGCTTGTGGCGCCATCTCTTCGATGCAACATTCCTTCGGCGGCGGTGCTC 5027  
DB |||||  
QY 5701 GCGGGGAGCTTGTGGCGCCATCTCTTCGATGCAACATTCCTTCGGCGGCGGTGCTC 5760  
DB |||||  
QY 5028 AACGGCTCAACCTACTACTTGGCTCTTCCCTAATGCAGAGTCGCATTAAGGAGAGCT 5087  
DB |||||  
QY 5761 AACGGCTCAACCTACTACTTGGCTCTTCCCTAATGCAGAGTCGCATTAAGGAGAGCT 5820  
DB |||||  
QY 5088 CGAGTATCTATGANTCGAAGTATGGGAATGGTATACCCGCAATTCCTCAGTGTCTTGAG 5147  
DB |||||  
QY 5821 CGAGTATCTATGANTCGAAGTATGGGAATGGTATACCCGCAATTCCTCAGTGTCTTGAG 5880  
DB |||||  
QY 5148 TCTCCTATCAGATATGCCCACTAAGCAACCGAGAGGAGATTCATGTAATTTTC 5207  
DB |||||  
QY 5881 TCTCCTATCAGATATGCCCACTAAGCAACCGAGAGGAGATTCATGTAATTTTC 5940  
DB |||||  
QY 5208 TCTGACTTTTGGTTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAA 5267  
DB |||||  
QY 5941 TCTGACTTTTGGTTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAA 6000  
DB |||||  
QY 5268 ATGTCCTTCTTGAGACAGTAATGAATGCCCAACCAATTAAGAAATCTTGTATCAGGA 5327  
DB |||||  
QY 6001 ATGTCCTTCTTGAGACAGTAATGAATGCCCAACCAATTAAGAAATCTTGTATCAGGA 6060  
DB |||||  
QY 5328 ACAAATCTTGTTCGAACTTTTCGGTGCCTTGAACCTATAAAATGTAGAGTGAATG 5387  
DB |||||  
QY 6061 ACAAATCTTGTTCGAACTTTTCGGTGCCTTGAACCTATAAAATGTAGAGTGAATG 6120  
DB |||||  
QY 5388 TCGGTAGGAATGGACGGGCAAACTCTTACCTTTGACCTTCAAGAGGTATGAGGT 5447  
DB |||||  
QY 6121 TCGGTAGGAATGGACGGGCAAACTCTTACCTTTGACCTTCAAGAGGTATGAGGT 6180  
DB |||||  
QY 5448 TTCTAGATCTGATGCACTTCAAGTCAAACTTTCAGTCAAACTTCAACCACTTCCGAA 5507  
DB |||||  
QY 6181 TTGTAGATCTGATGCACTTCAAGTCAAACTTTCAGTCAAACTTCAACCACTTCCGAA 6240  
DB |||||  
QY 5508 TCAGAGAAATCAAAGTTGTTGTCTACTATTGATCCAGCCAGTCGGTCTTGAACCTG 5567  
DB |||||  
QY 6241 TCAGAGAAATCAAAGTTGTTGTCTACTATTGATCCAGCCAGTCGGTCTTGAACCTG 6300  
DB |||||  
QY 5568 ACAAATAGTGTCTGTTTTCAGGTTCATCTTGTATGAATAAATCTAGTCTTGTATCTA 5627  
DB |||||  
QY 6301 ACAAATAGTGTCTGTTTTCAGGTTCATCTTGTATGAATAAATCTAGTCTTGTATCTA 6360  
DB |||||  
QY 5628 AATAATCTTTCAGAGCAAGGCGATATAATACCAAACTCTTAAACTCTTTTAAACCTTAA 5687  
DB |||||  
QY 6361 AATAATCTTTCAGAGCAAGGCGATATAATACCAAACTCTTAAACTCTTTTAAACCTTAA 6420  
DB |||||  
QY 5688 AAGGACAAATGATGTCCTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCATCA 5747  
DB |||||  
QY 6421 AAGGACAAATGATGTCCTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCATCA 6480  
DB |||||  
QY 5748 ACTTGAGGGGCACTATCTTGTCTTACAGAAATTTGGGAGATCGGATATCGAGAAAAGG 5807  
DB |||||  
QY 6481 ACTTGAGGGGCACTATCTTGTCTTACAGAAATTTGGGAGATCGGATATCGAGAAAAGG 5540  
DB |||||  
QY 5808 TACGCTGATTTTAAACGTGAAATTTATCTCAAGATCTTGCCTCGCGGTTTCGGTGATG 5867  
DB |||||  
QY 6541 TACGCTGATTTTAAACGTGAAATTTATCTCAAGATCTTGCCTCGCGGTTTCGGTGATG 6600  
DB |||||  
QY 5868 ACGGTGAATAACCTCTGACACATGCACCTCCCGGAGAGCGTCAACGCTTCTCTGTAAGCGG 5927  
DB |||||





1261 G C C C C C G G G A G A G A T A A A T T G T G C T T C C C G G T G T C A C C C C C A G C A G T G C A C G G A G 1320  
 1258 C A G T C C C A G G G G T G A A C C T C A G C C C T G C G A A G C A C A T C A T G A G A G A T C C A A G C 1317  
 1321 A G A G T T G C T G T T T G A T G A C A G T G T C C G G G A T C C C G T G T C C C A C C C A T G G C 1380  
 1318 C G C G C G A T G A C A T G A T G A T G A C ----- 1344  
 1381 A T C G A A A C A C T C A G A A G A A T G C C T T A A C T A G T G C G T A G A T T C C C T A G G 1440  
 1345 ----- G A C G A C T A A G A A T T C C C T A G G 1366  
 1441 C G C C G C G A A T T A A T T C G C C T T A G A C A T G A C T G T C C T C A G T T C A A G T T G G G C A C T T A G 1500  
 1367 C G C C G C G A A T T A A T T C G C C T T A G A C A T G A C T G T T C C T C A G T T C A A G T T G G G C A C T T A G 1426  
 1501 A G A A C C G G T T G C T A G A T T C T A A T C A A G A G A T G T C A G A A T G C C A T T G C C T A G A G 1560  
 1427 A G A G A C G G T C T G C T A G A T T C T A A T C A A G A G A T G T C A G A A T G C C A T T G C C T A G A G 1486  
 1561 A T G C A G C C T C A T T T T T G A T A C T T T T A T T T G T A A C C T A T A T A G T A G A T T T T T T 1620  
 1487 A T G C A G C C T C A T T T T T G A T A C T T T T A T T T G T A A C C T A T A T A G T A G A T T T T T T 1546  
 1621 G T C A T T T G T T C T T C T A C A G A G T T G C T C C T G A T C A G C C T A T C T G C A G C T G A T G A A 1680  
 1547 G T C A T T T G T T C T T C T G T A C G A C T T G C T C C T G A T C A G C C T A T C T G C A G C T G A T G A A 1606  
 1681 T A T C T T G T G A G G G T T G G A A A A T C A T T C G A G T T G A T G T T T T C T T G G T A T T C C C 1740  
 1607 T A T C T T G T G A G G G T T G G A A A A T C A T T C G A G T T G A T G T T T T C T T G T A T T T C C C 1666  
 1741 A C T C C T C T C A G A G A G A T T A A G T G A G A A G T T C G T T G T G C A A G C T T A T C A T A A 1800  
 1667 A C T C C T C T C A G A G A G A T T A A G T G A G A A G T T C G T T G T G C A A G C T T A T C A T A A 1726  
 1801 G C T T T A A T C G G T A G T T A T C A C A G T T A A A T G T A A C G A G T C A G C A C C G T G T A G A A 1860  
 1727 G C T T T A A T C G G T A G T T A T C A C A G T T A A A T G T A A C G A G T C A G C A C C G T G T A G A A 1786  
 1861 A T C T A A C A A T C G G T C A T C G T C A C C C G C C A C C C T G A T G C T G T A G C A T A G G 1920  
 1787 A T C T A A C A A T C G G T C A T C G T C C C G G C A C C G T C A C C C T G A T G C T G T A G C A T A G G 1846  
 1921 C T T G G T A T G C G G T A C T G C C G G C C T T T G C G G A T A T C G T C C A T T C C G A C A G A T C G C 1980  
 1847 C T T G G T A T G C G G T A C T G C C G G C C T T T G C G G A T A T C G T C C A T T C C G A C A G A T C G C 1906  
 1981 C A G T C A T A T G C G G T G T C T A G C G T A T A T G C T T G A T G C A A T T T C A T G C C A C C C G T 2040  
 1907 C A G T C A T A T G C G G T G T C T A G C G T A T A T G C T T G A T G C A A T T T C A T G C C A C C C G T 1966  
 2041 T C T C G A C A C T G T C C G A C C G T T T G C C G C C C C A G T C C T G C T C G C T T C G T A C T T G G 2100  
 1967 T C T C G A C A C T G T C C G A C C G T T T G C C G C C C C A G T C C T G C T C G C T T C G C T A C T T G G 2026  
 2101 A G C C A T A T C A G A C T A C G A T C A T G G G A C A C A C C C G T C T G T G A T C T A T G A A T C T A 2160  
 2027 A G C C A T A T C A G A C T A C G A T C A T G G G A C A C A C C C G T C T G T G A T C T A T G A A T C T A 2086  
 2161 A A T G T A A G T T A A A T C T A A A T A A A A A T A A G T C C C A G T T T C C A T A G A A C C T T A A 2220  
 2087 A A T G T A A G T T A A A T C T A A A T A A A A A T A A A T A A A T A A A T A A A T A A A T A A A T A A A 2146  
 2221 C A G A A T C G G T G A G A C T A G A C C T T C A A C A G A G A G A T C C A T C A C T G C T T G G C C A A 2280  
 2147 C A G A A T C G G T G A G A C T A G A C C T T C A A C A G A G A G A T C C A T C A C T G C T T G G C C A A 2206  
 2281 T A T G T T T C A G T C C C T C A G A G T A G T C T G T G A G T A G A A C T T C T G A A G G T T C A G 2340  
 2207 T A T G T T T C A G T C C C T C A G A G T T A G C T T T G T G A G T A G A A C T T C T G A G A G G C C T T A G A A A C G T 2266

2341 T G T T A A C T C C G C T G T A T T G A C G G C A T A T C C G T A C G T T G C A A A G T G T G T T G T A C C G G 2400  
 2267 T G T T A A C T C C G C T G T A T T G A C G G C A T A T C C G T A C G T T G C A A A G T G T G T T G T A C C G G 2326  
 2401 A G G A G T A A T T C C C A A A C T C T C T G A G A G T A G G A C C A A C A A A C A C A G A T C C A G C G T T 2460  
 2327 A G G A G T A A T T C C C A A A C T C T C T G A G A G T A G G A C C A A C A A A C A C A G A T C C A G C G T T 2386  
 2461 G T A C T T G A T C A A A T A A A A G A A C A T T C T C A T T T G C A G A T C A A G T T C A G A G C G T 2520  
 2387 G T A C T T G A T C A A A T A A A A G A A C A T T C T C A T T T G C A G A T C A A G T T C A G A G C G T 2446  
 2521 A C T G A T T G A C A T T T C C A A A G C C T G C T A G T A G T T G C A A C C G A T A G G T T G T A G A T G T G 2580  
 2447 A C T G A T T G A C A T T T C C A A A G C C T G C T A G T A G T T G C A A C C G A T A G G T T G T A G A T G T G 2506  
 2581 C A A T A C A C T T C G G T A C A A T T T C A A C C C T T G C A A C T G C A C A G C T T G T T G T G A A C A G C A T 2640  
 2507 C A A T A C A C T T C G G T A C A A T T T C A A C C C T T G C A A C T G C A C A G C T T G T T G T G A A C A G C A T 2566  
 2641 C T T C A A T T C T G C A A G C T C C T T G T C A T A T C G A C A G C C A A C A G A A T C A C C T G G A A T 2700  
 2567 C T T C A A T T C T G C A A G C T C C T T G T C A T A T C G A C A G C C A A C A G A A T C A C C T G G A A T 2626  
 2701 C A A T A C C A T G T T C A G C T T G A C A G A A G T C T G A G G C A A C G A A T C T G G A T C G A C G C T A T T 2760  
 2627 C A A T A C C A T G T T C A G C T T G A C A G A A G T C T G A G G C A A C G A A T C T G G A T C A C G C T A T T 2686  
 2761 T A T C A G C A A T A A C T A G A A C T T C A G A A G C C C A G A G C A T G T C A A T A C T A C A C A G G C T G 2820  
 2687 T A T C A G C A A T A A C T A G A A C T T C A G A A G C C C A G A G C A T G T C A A T A C T A C A C A G G C T G 2746  
 2821 A F T G T C A T T T T G A A C A T C A T C T T G C A G C A G T A A C G A A C T G T T C C T G G A C A A T A 2880  
 2747 A F T G T C A T T T T G A A C A T C A T C T T G C A G C A G T A A C G A A C T G T T C C T G G A C A A T A 2806  
 2881 T T T T G T C A C A C T A G A A C A G T T T C T G T T C G T A A G C C A T A G C A G C T A C T C C C T G G G C G 2940  
 2807 T T T T G T C A C A C T A G A A C A G T T T C T G T T C G T A A G C C A T A G C A G C T A C T G C C T G G G C G 2866  
 2941 C T C T G C T A G C A C A T A C A C T T A G C A C C A C C T T T G T G G G C A A C G T A G A T G A T T C T G G G G 3000  
 2867 C T C T G C T A G C A C A T A C A C T T A G C A C C A C C T T T G T G G G C A A C G T A G A T G A T T C T G G G G 2926  
 3001 T A A G G T A C C A T C C T T C A A T A G T C T T G A G A T G C A A A A C A A T T T C T T G C A A C C A G C A C T T 3060  
 2927 T A A G G T A C C A T C C T T C T T A G T G A G A T G C A A A A C A A T T T C T T G C A A C C A G C A C T T 2986  
 3061 T G G C A G G A A C C C C A C A T C A G G A A G T G G A A G C A G A A T T G C G G T T C C A C C A G A A T A T 3120  
 2987 T G C A G G A A C C C C A C A T C A G G A A G T G G A A G C A G A A T T G C G G T T C C A C C A G A A T A T 3046  
 3121 A G A G C C A A C T T T C T C A A T A G T C T T G C A A A C C A G A C A G C A C T A C C A G G C A A G C T 3180  
 3047 A G A G C C A A C T T T C A A T A G T C T T G C A A A C C A G A C A G C A C T A C A C C A G G C A A G C T 3106  
 3181 C A A C T T G C A A C G T C T C C G T A G T T A G A C T T C A T G A A T T T C C T G A C G T T A T C T A T A G A 3240  
 3107 C A A C T T G C A A C G T C T C C G T A G T T A G A C T T C A T G A A T T T C C T G A C G T T A T C T A T A G A 3166  
 3241 G A T C A A T G G C T C T T A A C G T T A T C T G C A A T T G A A G T T C C T C T G G A A A G A G C T T 3300  
 3167 G A T C A A T G G C T C T T A A C G T T A T C T G C A A T T G C A A T A A G T T C C T C T G G A A A G A G C T T 3226  
 3301 C T A A C A C A G T C T C T T C A A A G C A C T C C A T C A A A C T T G C C A G T T A G T T C T A A A A G G G C T T 3360  
 3227 C T A A C A C A G T C T C T T C A A A G C A C T C C A T C A A C T T G C C A G T T A G T T C T A A A A G G G C T T 3286  
 3361 T G T C A C C A T T T T G A C A A C A T T G T C G A A A T T G T T G A C T A A T T C C A T A T C T G T C G G 3420  
 3287 T G T C A C C A T T T T G A C A A C A T T G T C G A A A T T G T T G A C T A A T T C C A T A T C T G T C G G 3346  
 3421 T T T T C T G T A G A C A C A G A A G G C A T C T T C A A T T T T T G T G A G A G G C C T T A G A A A C G T 3480

Db 3347 TTTTCTGGATAGGACGAGAGGGGATCTCTCAATTTCTTGAGAGGCGCTTGAAGACGT 3406  
Qy CAATTTTGCACAATTAATAGACCTTCAGAAAGGAGCTTCTTTAGTTTGGATTCTTCTT 3540  
Db CAATTTTGCACAATTAATAGACCTTCAGAAAGGAGCTTCTTTAGTTTGGATTCTTCTT 3466  
Qy TAGTTTGTCTCTGGTGTATCTGGCTTGGATCTCTCTTCTCTCTAGTACCTTTAGGG 3600  
Db TAGTTTGTCTCTGGTGTATCTGGCTTGGATCTCTCTTCTCTCTAGTACCTTTAGGG 3526  
Qy ACTTCATATCCAGGTTTCTCTCCACCTCGTCCAAAGCTCACACCGTACTTGGACATCTAA 3660  
Db ACTTCATATCCAGGTTTCTCTCCACCTCGTCCAAAGCTCACACCGTACTTGGACATCTAA 3586  
Qy CTAATGCAAAATAAATAAATAGTCAGACATCTCCAGGCTATATCTCTCTGGATTAGCTT 3720  
Db CTAATGCAAAATAAATAAATAGTCAGACATCTCCAGGCTATATCTCTCTGGATTAGCTT 3646  
Qy CTGCAAGTTTCATCAGTTCCTCCCTTAATTTTAGGGTTCAACAAAACCTTCGTCGTAATA 3780  
Db CTGCAAGTTTCATCAGTTCCTCCCTTAATTTTAGGGTTCAACAAAACCTTCGTCGTAATA 3706  
Qy ACCGTTTGGTATAGAACTTCTGGAGCATTTCTTACGATCCCAAGGTGCTTCCA 3840  
Db ACCGTTTGGTATAGAACTTCTGGAGCATTTCTTACGATCCCAAGGTGCTTCCA 3766  
Qy TGGCTCTAAGACCTTTGATTGGCCAAACAGGAAGTCGGTTCCAAAGTGACAGAAACCAA 3900  
Db TGGCTCTAAGACCTTTGATTGGCCAAACAGGAAGTCGGTTCCAAAGTGACAGAAACCAA 3826  
Qy CACTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCAATCCAAATTCGATACCCA 3960  
Db CACTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCAATCCAAATTCGATACCCA 3886  
Qy GCACTTTTGTAGTTGCTCCAGATGACACCTTTATACCAAAACCGTGACGAGATT 4020  
Db GCACTTTTGTAGTTGCTCCAGATGACACCTTTATACCAAAACCGTGACGAGATT 3946  
Qy GGTAGATCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4080  
Db GGTAGATCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4006  
Qy GGCCCAACGATTAATAGAAAGTCAGCCACCAAGTAGTAATAGACCATCGGGGCGGT 4140  
Db GGCCCAACGATTAATAGAAAGTCAGCCACCAAGTAGTAATAGACCATCGGGGCGGT 4066  
Qy CAGTAGTCABAGACGCCAACAAATTTTCACTGACAGGAACTTTTTCGACATCTTCAGAA 4200  
Db CAGTAGTCABAGACGCCAACAAATTTTCACTGACAGGAACTTTTTCGACATCTTCAGAA 4126  
Qy GTTCGTATTCAGTAGTCAATTCGCGAGCATCAATAATGGGGATTATACCAAGAACACAG 4260  
Db GTTCGTATTCAGTAGTCAATTCGCGAGCATCAATAATGGGGATTATACCAAGAACACAG 4186  
Qy TGGAGTCACATCTACCAACTTTGGGTCTCAGAAAAGGATTAACAGTTCTACTACCGC 4320  
Db TGGAGTCACATCTACCAACTTTGGGTCTCAGAAAAGGATTAACAGTTCTACTACCGC 4246  
Qy CATTAGTCAGAACTTTTCAATTCGCGGCTCAGAAAAGGATTAACAGTTCTACTACCGC 4380  
Db CATTAGTCAGAACTTTTCAATTCGCGGCTCAGAAAAGGATTAACAGTTCTACTACCGC 4306  
Qy TAGCGGGCAAGATGCAACTTTATCAACAGGGTCTTATAGATAACCCCTAGCGCTGGGA 4440  
Db TAGCGGGCAAGATGCAACTTTATCAACAGGGTCTTATAGATAACCCCTAGCGCTGGGA 4366  
Qy TCATCTTTGGACAACTCTTTCTGCCAAATCTAGTCCAAATCACTTCATTTGATACCAT 4500  
Db TCATCTTTGGACAACTCTTTCTGCCAAATCTAGTCCAAATCACTTCATTTGATACCAT 4426  
Qy TATTGTACAACTTGAGCAAGTTGTGATCAGTCTCTCAATTTGGTCTCTGTACAGGATG 4560

Db 4427 TATTGTACAACTTGAGCAAGTTGTGATCAGCTCTCAAAATTTGGTCTCTGTAAACGATG 4486  
Qy ACTCAAACTTGCACATTAACCTTGAAGCTCAGTCGATTGAGTGAACCTTGATCAGGTTGTGA 4620  
Db ACTCAAACTTGCACATTAACCTTGAAGCTCAGTCGATTGAGTGAACCTTGATCAGGTTGTGA 4546  
Qy GCTGGTCAGCAGCATAGGGAACACCGCTTTTCTTACCAAACTCAAGGAATATCAAACT 4680  
Db GCTGGTCAGCAGCATAGGGAACACCGCTTTTCTTACCAAACTCAAGGAATATCAAACT 4606  
Qy CTGCAACACTTGGGTATGAGGATAGCAAGGAAATGTCTACTTGAAGTCGGACAGTGTAG 4740  
Db CTGCAACACTTGGGTATGAGGATAGCAAGGAAATGTCTACTTGAAGTCGGACAGTGTAG 4666  
Qy TGTAGTCTTGCAGAAATCTGAAGCCGTATTTTATATCAGTGAAGTCATCATTGAGTCGAGAGA 4800  
Db TGTAGTCTTGCAGAAATCTGAAGCCGTATTTTATATCAGTGAAGTCATCATTGAGTCGAGAGA 4726  
Qy TCCTCTACGCGGACGACATCGTGGCGGA 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4786  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4846  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4906  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4966  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5026  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5086  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5146  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5206  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5266  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5326  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5386  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5446  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5506  
Qy TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 4828  
Db TCCTCTACGCGGACGACATCGTGGCGGACCTGACGGGGGGGGGGCGCTGAGGTCTGC 5566

QY 4829 ----- 4828  
Db 5567 CCGGGGATCGCAGTGGTGAATCAACCATGCATCATCAGGAGTACGGATAAAATGCTTGAT 5626  
QY 4829 ----- 4828  
Db 5627 GGTGGAGAGGAGCAATAAATTCGGTCAGCGAGTTTAGTCTGACCATCTCATCTGTAACATC 5686  
QY 4829 ----- 4828  
Db 5687 ATTGGCAACGCTACCTTTGGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCAT 5746  
QY 4829 ----- 4828  
Db 5747 CAATCGATAGATTGTGCGACCTGATTGCCGACATTATCGGAGGCCATTTATACCCATA 5806  
QY 4829 ----- 4828  
Db 5807 TAAATCAGCATCCATGTTGGAAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5866  
QY 4829 ----- 4828  
Db 5867 ATGGCTATAACACCCCTTGATTTACTGTTTATGTAACGACACAGTTTATTGTTATGA 5926  
QY 4829 ----- 4828  
Db 5927 TGATATATTTTATCTTGTGCAATGTAACATCAGAGATTTTGGAGACAAACGTGGCTTTC 5986  
QY 4829 ----- CTTGCAAGTCCGATCACCGGCGCACAGGTGCGGTTGCTGGCGCTATAT 4879  
Db 5987 CCCCCCCCCCTCGAGTCCGATCACCGGCGCACAGGTGCGGTTGCTGGCGCTATAT 6046  
QY 4880 CGCGAGATCAACGATGGGGAAGATCGGGCTCGGCACTTGGGCTCATAGAGCGCTGTTTT 4939  
Db 6047 CGCCGACATCACCGATGGGGAAGATCGGGCTCGGCACTTGGGCTCATAGAGCGCTGTTTT 6106  
QY 4940 CGGCTGGGTATGTTGGCAGGCGCGTGGCGGGGACTGTTGGGCGCCATCTCTTGCA 4999  
Db 6107 CGGCTGGGTATGTTGGCAGGCGCGTGGCGGGGACTGTTGGGCGCCATCTCTTGCA 6166  
QY 5000 TGACCATTTCTTGGCGGGGCTGCTCAACGGCTCAACCTACTACTGGGCTGCTTCCCT 5059  
Db 6167 TGACCATTTCTTGGCGGGGCTGCTCAACGGCTCAACCTACTACTGGGCTGCTTCCCT 6226  
QY 5060 AATGACAGAGTCGCATAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGGAATGGT 5119  
Db 6227 AATGACAGAGTCGCATAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGGAATGGT 6286  
QY 5120 GATACCGCATTTCTCAGTGTCTTGAAGTCTCTATCAGATTAATGCGCAACTAAAGCAAC 5179  
Db 6287 GATACCGCATTTCTCAGTGTCTTGAAGTCTCTATCAGATTAATGCGCAACTAAAGCAAC 6346  
QY 5180 CGGAGGAGGAGATTTCTATGTAATAATTTCTCTGACTTTTGGTCTATCAGTATGCGCAACTG 5239  
Db 6347 CGGAGGAGGAGATTTCTATGTAATAATTTCTCTGACTTTTGGTCTATCAGTATGCGCAACTG 6406  
QY 5240 TGAGACTATCTCGGTTATGACGAGAAATTTCTCTTGGTCTATCAGTATGCGCAACTG 5299  
Db 6407 TGAGACTATCTCGGTTATGACGAGAAATTTCTCTTGGTCTATCAGTATGCGCAACTG 6466  
QY 5300 ACCAATAAGAAATCTTGTATCAGGAACAACTTCTTGTTCGAATTTTTCGGTGCC 5359  
Db 6467 ACCAATAAGAAATCTTGTATCAGGAACAACTTCTTGTTCGAATTTTTCGGTGCC 6526  
QY 5360 TTGAATATAAATAGATGATGATGTTGGTATGGAATGGAGCGGCAAAATGCTTACC 5419  
Db 6527 TTGAATATAAATAGATGATGATGTTGGTATGGAATGGAGCGGCAAAATGCTTACC 6586  
QY 5420 TTCTGGACCTTCAAGAGATGATGATGTTGGTATGGAATGGAGCGGCAAAATGCTTACC 5479  
Db 6587 TTCTGGACCTTCAAGAGATGATGATGTTGGTATGGAATGGAGCGGCAAAATGCTTACC 6646

QY 5480 GTTGTATTTTCTGTTCAACCATTCGGAATCCAGAGAAATCAAAAGTTGTTGTCTACTATT 5539  
Db 6647 GTTGTATTTTCTGTTCAACCATTCGGAATCCAGAGAAATCAAAAGTTGTTGTCTACTATT 6706  
QY 5540 GATCCAAAGCCAGTGGGCTTTGAAACTGACAATAGTGTGCTCGTGTGTTGAGTCTATCTT 5599  
Db 6707 GATCCAAAGCCAGTGGGCTTTGAAACTGACAATAGTGTGCTCGTGTGTTGAGTCTATCTT 6766  
QY 5600 TGTATGATAAATCTAGTCTTTGATCTAAATATCTTGRGRCGCCAGGCGGATAAATACC 5659  
Db 6767 TGTATGATAAATCTAGTCTTTGATCTAAATATCTTGRGRCGCCAGGCGGATAAATACC 6826  
QY 5660 CAAATCTAAAACTCTTTTAAAAAGCTTAAAAAGGCAAGTATGCTGCTGTTTAAACCCC 5719  
Db 6827 CAAATCTAAAACTCTTTTAAAAAGCTTAAAAAGGCAAGTATGCTGCTGTTTAAACCCC 6886  
QY 5720 AAATCAGCTCTAGTCTGATCTCTCATCAACTTGAAGGGCACTATCTGTTTGAAGAAAT 5779  
Db 6887 AAATCAGCTCTAGTCTGATCTCTCATCAACTTGAAGGGCACTATCTGTTTGAAGAAAT 6946  
QY 5780 TTGCGGAGATCGGATATCGAGAAAAAGGTACGCTGATTTTAAAAAGTGAATTTATCTCAA 5839  
Db 6947 TTGCGGAGATCGGATATCGAGAAAAAGGTACGCTGATTTTAAAAAGTGAATTTATCTCAA 7006  
QY 5840 GATCTCTGCTCGCGGCTTTCGGTGATGAGGTGAAAAACCTCTGACACATGAGCTCCCG 5899  
Db 7007 GATCTCTGCTCGCGGCTTTCGGTGATGAGGTGAAAAACCTCTGACACATGAGCTCCCG 7066  
QY 5900 GAGACGCTCAAGCTTGTCTGTAAGCGGATCCGGGAGCAGACAAAGCCGCTCAGGGCGCG 5959  
Db 7067 GAGACGCTCAAGCTTGTCTGTAAGCGGATCCGGGAGCAGACAAAGCCGCTCAGGGCGCG 7126  
QY 5960 TCAGCGGCTTTCGGGCTGTCGGGCGCAGCAATGACCCAGTCACTGAGCAGTAGGGA 6019  
Db 7127 TCAGCGGCTTTCGGGCTGTCGGGCGCAGCAATGACCCAGTCACTGAGCAGTAGGGA 7186  
QY 6020 GTGTATATCTGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATGC 6079  
Db 7187 GTGTATATCTGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATGC 7246  
QY 6080 GGTGTAAATACCCACAGATGCGTAAGGAGAAATACCGCATCAGCGCTCTTCGCTT 6139  
Db 7247 GGTGTAAATACCCACAGATGCGTAAGGAGAAATACCGCATCAGCGCTCTTCGCTT 7306  
QY 6140 CCTGCTCACTGACTCGCTCGGCTCGGCTCGGCTGCGGCGAGCGGTATCAGCTCACT 6199  
Db 7307 CCTGCTCACTGACTCGCTCGGCTCGGCTCGGCTGCGGCGAGCGGTATCAGCTCACT 7366  
QY 6200 CAAAGCGGTAATACGTTATCCAGAAATCAGGGAATACCGAGAAAGACATGTGAG 6259  
Db 7367 CAAAGCGGTAATACGTTATCCAGAAATCAGGGAATACCGAGAAAGACATGTGAG 7426  
QY 6260 CAAAGCGGTAATACGTTATCCAGAAATCAGGGAATACCGAGAAAGACATGTGAG 6319  
Db 7427 CAAAGCGGTAATACGTTATCCAGAAATCAGGGAATACCGAGAAAGACATGTGAG 7486  
QY 6320 GGTCTCGGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGGTGGCG 6379  
Db 7487 GGTCTCGGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGGTGGCG 7546  
QY 6380 CGACGAGCTATATAAGATACGAGCGTTCCTCCCTGGAAGCTCCCTGCGGCTCTCTG 6439  
Db 7547 CGACGAGCTATATAAGATACGAGCGTTCCTCCCTGGAAGCTCCCTGCGGCTCTCTG 7606  
QY 6440 TTCGAGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGGTGGCG 6499  
Db 7607 TTCGAGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGGTGGCG 7666  
QY 6500 TTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAGCTGG 6559  
Db 7667 TTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAGCTGG 7726  
QY 6560 GCTGTGTGACGAACCCCCCGTTACGCGGACCGCTTCGCGCTTATCCGGTAACTATCTGTC 6619

7727	Db		GCTGTGTGCAGAAACCCCGCGTTTCAGCCCGACCGCTCGCGCTTATCCGGTAACTATCGTC	7786
6620	Qy		TTGAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA	6679
7787	Db		TTGAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA	7846
6680	Qy		TTAGCAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGCCCTAACTACG	6739
7847	Db		TTAGCAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGCCCTAACTACG	7906
6740	Qy		GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTTACCTTCGGAA	6799
7907	Db		GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTTACCTTCGGAA	7966
6800	Qy		AAAGAGTTGGTAGTCTTTGATTCGGCAAAACAAACCCCGCTGGTAGCGGTGGTTTTTG	6859
7967	Db		AAAGAGTTGGTAGTCTTTGATTCGGCAAAACAAACCCCGCTGGTAGCGGTGGTTTTTG	8026
6860	-Qy		TTTCCACGACGACGATTACGGCGAGAAAAGGATCTCAAGAGATCCCTTTGATCTCTTT	6919
8027	Db		TTTCCACGACGACGATTACGGCGAGAAAAGGATCTCAAGAGATCCCTTTGATCTCTTT	8086
6920	Qy		CTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGAT	6979
8087	Db		CTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGAT	8146
6980	Qy		TATCAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCT	7039
8147	Db		TATCAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCT	8206
7040	Qy		AAAGTATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCAGCTA	7099
8207	Db		AAAGTATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCAGCTA	8266
7100	Qy		TCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTTGCCTGACTCCCGTCTGTAGATAA	7159
8267	Db		TCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTTGCCTGACTCCCGTCTGTAGATAA	8326
7160	Qy		CTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATCATACCGGAGACCCAC	7219
8327	Db		CTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATCATACCGGAGACCCAC	8386
7220	Qy		GCTCACCGGCTCCAGATTATCAGCAATAAACCCAGCCGCGAAGGCGCGAGCGCAGAA	7279
8387	Db		GCTCACCGGCTCCAGATTATCAGCAATAAACCCAGCCGCGAAGGCGCGAGCGCAGAA	8446
7280	Qy		GTGGTCTGCAACTTATCCCGCTCCATCAGTCTATTAAATGTTGTCGGGGAAGCTAGAG	7339
8447	Db		GTGGTCTGCAACTTATCCCGCTCCATCAGTCTATTAAATGTTGTCGGGGAAGCTAGAG	8506
7340	Qy		TAAGTAGTTCCGCAAGTTAATAGTTTGCACACGTTTGTCCATATGCTGCAGGCAATCGTG	7399
8507	Db		TAAGTAGTTCCGCAAGTTAATAGTTTGCACACGTTTGTCCATATGCTGCAGGCAATCGTG	8566
7400	Qy		TGTCAACGCTCGTGTGGTATGCGTTCAATTCAGTCCGGTTCGCAACGATCAAGGCGAG	7459
8567	Db		TGTCAACGCTCGTGTGGTATGCGTTCAATTCAGTCCGGTTCGCAACGATCAAGGCGAG	8626
7460	Qy		TTACATGATCCCCATGTTGCAAAAAGCGGTAGTCTCTTCGGTCTCCGATCGTTG	7519
8627	Db		TTACATGATCCCCATGTTGCAAAAAGCGGTAGTCTCTTCGGTCTCCGATCGTTG	8686
7520	Qy		TCAGAAGTAAGTTGCCCGCAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTC	7579
8687	Db		TCAGAAGTAAGTTGCCCGCAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTC	8746
7580	Qy		TTACTGTATGCCATCCGTAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAT	7639
8747	Db		TTACTGTATGCCATCCGTAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAT	8806
7640	Qy		TCTGAGAAATAGTGTATGGCGGACCGAGTTGCTCTTTCGCCCGCGTCAACACGGGATAATA	7699

Db	8867	TCTGAGAAATAGTGATCATCGCGGACCGAGTTGCTCTTCCCGCGCGTCAACACGGGATPATA	8866
Qy	7700	CCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGAAAAACGTTCTTCGGGGCGAA	7759
Db	8867	CCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGAAAAACGTTCTTCGGGGCGAA	8926
Qy	7760	AACCTCTCAAGGATCTTACCGCTGTTGAGATCCGAGTTGATGTAACCCACTCGTCACCCCA	7819
Db	8927	AACCTCTCAAGGATCTTACCGCTGTTGAGATCCGAGTTGATGTAACCCACTCGTCACCCCA	8986
Qy	7820	ACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC	7879
Db	8987	ACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC	9046
Qy	7880	AAAATGCCGCAAAAAGGGAATAGGGCGACACGGAATGTTGAACTCATACTCTTCC	7939
Db	9047	AAAATGCCGCAAAAAGGGAATAGGGCGACACGGAATGTTGAACTCATACTCTTCC	9106
Qy	7940	TTTTTTCATATTTATGGAAGCATTTTATCAGGGTTATTGCTCATGAGCGGATACATATTG	7999
Db	9107	TTTTTTCATATTTATGGAAGCATTTTATCAGGGTTATTGCTCATGAGCGGATACATATTG	9166
Qy	8000	AATGTATTTAGAAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGGAAAAGTGCAC	8059
Db	9167	AATGTATTTAGAAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGGAAAAGTGCAC	9236
Qy	8060	CTGAGCTCTAAGAAACCAATTATTCATGACATTAACTTAAATAAGCGGTATCACGA	8119
Db	9227	CTGAGCTCTAAGAAACCAATTATTCATGACATTAACTTAAATAAGCGGTATCACGA	9286
Qy	8120	GGCCCTTTGCTCTTCAAGAAATTAATCTCATGTTTGAAGCTTATCATGAAAGCTGAC	8179
Db	9287	GGCCCTTTGCTCTTCAAGAAATTAATCTCATGTTTGAAGCTTATCATGAAAGCTGAC	9346
Qy	8180	TCATGTTGGTATTTGTAATAGCGCAGATCGGGAACACTGAAAAATAACAGTTATTATT	8239
Db	9347	TCATGTTGGTATTTGTAATAGCGCAGATCGGGAACACTGAAAAATAACAGTTATTATT	9406
Qy	8240	CG 8241	
Db	9407	CG 9408	
RESULT 5			
AAAF83654			
ID	AAAF83654 standard; DNA; 9407 BP.		
XX	AAAF83654;		
AC			
XX			
DT	23-JUL-2001 (first entry)		
XX			
DE	pPIC9K-lunasin gene construct sequence.		
XX			
KW	Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene; ds.		
XX			
OS	Glycine max.		
XX			
FH	Key		
PD	CDS		
FT	Location/Qualifiers		
FT	1222..1353		
FT	/*tag= a		
FT	/note= "lunasin DNA sequence"		
XX			
PN	WO200134808-A2.		
XX			
PD	17-MAY-2001.		
XX			
PF	12-NOV-2000; 2000WO-US031211.		
XX			
PR	12-NOV-1999; 99US-0165334P.		
XX			
PA	(FILG-) FILGEN BIOSCIENCES INC.		
PI	Galvez AF;		



XX WPI; 2001-343605/36.  
DR P-PSDB; AAB62613.  
XX  
PT Large scale production of lunasin, a cancer preventive peptide from  
PT soybean, by preparing lunasin gene constructs using pichia expression  
PT vectors, optimizing lunasin expression parameters and purifying lunasin.  
XX  
PS Disclosure; Page 65-67; 67pp; English.  
XX  
CC The invention provides a method of producing lunasin, a cancer preventive  
CC peptide from soybean, or its active variant by recombinant DNA technology  
CC in large quantities. The method comprises preparing lunasin gene  
CC constructs or its variants using protein expression vectors, optimizing  
CC lunasin expression parameters for large-scale production and performing a  
CC series of isolation and purification to obtain large quantities of  
CC biologically active recombinant lunasin peptide. The lunasin peptide  
CC exhibits inhibitory effects against malignant transformation of cells  
CC induced by chemical carcinogens and viral oncogenes. The present sequence  
CC represents the pPIC9K-lunasin gene construct sequence  
XX  
SQ Sequence 9407 BP; 2533 A; 2232 C; 2117 G; 2525 T; 0 U; 0 Other;  
Query Match 81.0%; Score 6675.6; DB 4; Length 9407;  
Best Local Similarity 85.2%; Fred. No. 0;  
Matches 8082; Conservative 0; Mismatches 84; Indels 1316; Gaps 4;  
1 AGATCTAATCCAAAGACGAAAGGTTGAATGAAGAACTTTTGGCATCCGACATCCACAG 60  
1 AGATCTAATCCAAAGACGAAAGGTTGAATGAAGAACTTTTGGCATCCGACATCCACAG 60  
61 GTCATTTCTACATATAGTCCAAACGCAACAGAGGGGATACACTAGCAGACGACCGT 120  
61 GTCCATTTCTACATATAGTCCAAACGCAACAGAGGGGATACACTAGCAGACGACCGT 120  
121 TGCAACCGCAGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGCCATCCGAAACAC 180  
121 TGCAACCGCAGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGCCATCCGAAACAC 180  
181 AGCCGAGTTATGGCTTTGATGAGCTCGCTCATTTCCAAATCTCTTATAGGCTACTA 240  
181 AGCCGAGTTATGGCTTTGATGAGCTCGCTCATTTCCAAATCTCTTATAGGCTACTA 240  
241 ACACCATGACTTATATGCTTGTATCTCTGCTCCGCTCCGAGGTTTCATGTTTGTTA 300  
241 ACACCATGACTTATATGCTTGTATCTCTGCTCCGCTCCGAGGTTTCATGTTTGTTA 300  
301 TTTCCGAATGCAACAGCTCCGCTATTAACCCGCAACATCACTCCAGATGAGGCTTTCTG 360  
301 TTTCCGAATGCAACAGCTCCGCTATTAACCCGCAACATCACTCCAGATGAGGCTTTCTG 360  
361 AGTGTGGGTCAATAGTTTATGTTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420  
361 AGTGTGGGTCAATAGTTTATGTTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420  
421 GTCCTTGGAACTTATATGCAAAAAGCGTGATCTCATCCAGATGAACCTAAGTTTGGTTCG 480  
421 GTCCTTGGAACTTATATGCAAAAAGCGTGATCTCATCCAGATGAACCTAAGTTTGGTTCG 480  
481 TTGAATGCTAACCGCAGTTGGTTCAAAAGAACTTCCAAAAGTCCGCATACCGTTTGT 540  
481 TTGAATGCTAACCGCAGTTGGTTCAAAAGAACTTCCAAAAGTCCGCATACCGTTTGT 540  
541 CTGTGTTGGTATGATGACAAATGCTCAAAAATATCTCAATTAATGCTTAGCCAGTCT 600  
541 CTGTGTTGGTATGATGACAAATGCTCAAAAATATCTCAATTAATGCTTAGCCAGTCT 600  
601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGCAACGCAAAATGGGGAAACACCCGCT 660  
601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGCAACGCAAAATGGGGAAACACCCGCT 660  
661 TTTTGGAGTATGATGCTCCACATGTTGATGCTTCCAGATTTCTGGTGGGAATACT 720

DB 661 TTTTGGAGTATGATGCTCCACATGTTCTCCAGATTTCTGGTGGGAATACT 720  
QY 721 GCTGATAGCTAAACGTTTCATGATCAAAAATTTAACTGTTCTAACCCCTACTTGCAGCAAT 780  
DB 721 GCTGATAGCTAAACGTTTCATGATCAAAAATTTAACTGTTCTAACCCCTACTTGCAGCAAT 780  
QY 781 ATATATAACAGAGGAAGCTGCTCTTAACCTTTTATTAATCATCATTTATAGCTT 840  
DB 781 ATATATAACAGAGGAAGCTGCTCTTAACCTTTTATTAATCATCATTTATAGCTT 840  
QY 841 ACTTTTCATTAATTCGACTGTTTCCAAATTCGAAAGCTTTTGAATTTAAAGCTTTTAAACGA 900  
DB 841 ACTTTTCATTAATTCGACTGTTTCCAAATTCGAAAGCTTTTGAATTTAAAGCTTTTAAACGA 900  
QY 901 CAACTTGAGAGATCAAAAAACAATAATTTTGAAGATCCAAAAGATGATTTCT 960  
DB 901 CAACTTGAGAGATCAAAAAACAATAATTTTGAAGATCCAAAAGATGATTTCT 960  
QY 961 TCAATTTTACTGCGAGTTTATTCGAGCATCTCCCGCATTAGCTTCCAGTCAACACT 1020  
DB 961 TCAATTTTACTGCGAGTTTATTCGAGCATCTCCCGCATTAGCTTCCAGTCAACACT 1020  
QY 1021 ACAACAGAGATGAAACGCGCAAAATTCGGCTGAAAGCTGTCTCGGTTACTCAGATTTA 1080  
DB 1021 ACAACAGAGATGAAACGCGCAAAATTCGGCTGAAAGCTGTCTCGGTTACTCAGATTTA 1080  
QY 1081 GAAGGGATTCGATGTTGCTGTTTGGCATTTCCACAGCACAATAAAGGGTTATTG 1140  
DB 1081 GAAGGGATTCGATGTTGCTGTTTGGCATTTCCACAGCACAATAAAGGGTTATTG 1140  
QY 1141 TTTTAAATACTACTATTTCGAGCATCTCTGCTAAAAGAAAGGGGTATCTCTCGAGAA 1200  
DB 1141 TTTTAAATACTACTATTTCGAGCATCTCTGCTAAAAGAAAGGGGTATCTCTCGAGAA 1200  
QY 1201 AGAGAGGTGAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1260  
DB 1201 AGAGAGGTGAAGCTTACGTAT---CCAAATGGCAGCACCAGCAAGTAGCTGCGCAAG 1260  
QY 1261 GCGCGCGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
DB 1261 GCGCGCGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
QY 1321 AGAGAGGTGCTGTTTGTGATGACAGTGTCCGGGATTTCCGTTGGTGTCTTCCACCCATGSCC 1380  
DB 1321 AGAGAGGTGCTGTTTGTGATGACAGTGTCCGGGATTTCCGTTGGTGTCTTCCACCCATGSCC 1380  
QY 1344 GCGCGCGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1344  
DB 1344 GCGCGCGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1344  
QY 1381 ATCGAGAACACTCAAGAGAGAGATGTCCTTCTAACTAGTGGGTAGAAATTCCTTAGGG 1440  
DB 1381 ATCGAGAACACTCAAGAGAGAGATGTCCTTCTAACTAGTGGGTAGAAATTCCTTAGGG 1440  
QY 1441 CGGCGCGAAATTAATTCGCTTTAGACATGCTGTTTCTCAGTTCAAGTTGGGCACTTACG 1500  
DB 1441 CGGCGCGAAATTAATTCGCTTTAGACATGCTGTTTCTCAGTTCAAGTTGGGCACTTACG 1500  
QY 1501 AGAAGACCGCTTCTGCTAGATCTTAATCAAGAGGATGTCAGAAATCCCAATTTGCCCTGAGAG 1560  
DB 1501 AGAAGACCGCTTCTGCTAGATCTTAATCAAGAGGATGTCAGAAATCCCAATTTGCCCTGAGAG 1560  
QY 1561 ATGACAGGCTTCAATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
DB 1561 ATGACAGGCTTCAATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
QY 1621 GTCAATTTGTTTCTCTGATCGAGCTTCTGCTGATCAGGCTATCTCGAGCTGATGAA 1680  
DB 1621 GTCAATTTGTTTCTCTGATCGAGCTTCTGCTGATCAGGCTATCTCGAGCTGATGAA 1680  
QY 1681 TATCTTGTGGTAGGGGTTTGGGAAATCATTTGAGTGTGTTTCTTGGTATTTCC 1740  
DB 1681 TATCTTGTGGTAGGGGTTTGGGAAATCATTTGAGTGTGTTTCTTGGTATTTCC 1740  
QY 1741 ACTCTCTTTCAGAGTACAGAGATTAAGTGAAGCTTCTGTTTGGTGAAGCTTATCGATAA 1800  
DB 1741 ACTCTCTTTCAGAGTACAGAGATTAAGTGAAGCTTCTGTTTGGTGAAGCTTATCGATAA 1800  
QY 1767 ACTCTCTTTCAGAGTACAGAGATTAAGTGAAGCTTCTGTTTGGTGAAGCTTATCGATAA 1726  
DB 1767 ACTCTCTTTCAGAGTACAGAGATTAAGTGAAGCTTCTGTTTGGTGAAGCTTATCGATAA 1726

QY 1801 GCTTTAATCGGCTAGTTTATACAGTTAAATTTGCTAAACGACGTACAGGACCGGTGATGAA 1860  
DB 1727 GCTTTAATCGGCTAGTTTATACAGTTAAATTTGCTAAACGACGTACAGGACCGGTGATGAA 1786  
QY 1861 ATCTAAACAATGCGCTCATCGTCTCGGCAACCGTCAACCGCTGAGTGTGATAGGCAATAGG 1920  
DB 1787 ATCTAAACAATGCGCTCATCGTCTCGGCAACCGTCAACCGCTGAGTGTGATAGGCAATAGG 1846  
QY 1921 CTTGGTTATGCGGCTACTGCGGCGCTCTTGGCGGCGATATCGTCCATTCGACAGCATGCG 1980  
DB 1847 CTTGGTTATGCGGCTACTGCGGCGCTCTTGGCGGCGATATCGTCCATTCGACAGCATGCG 1906  
QY 1981 CAGTCACTATGCGGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT 2040  
DB 1907 CAGTCACTATGCGGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT 1966  
QY 2041 TCTCGAGACACTGTCCGACCGCTTTGGCGGCGCGCGAGTCCGTGCTTGGCTTGGCTACTTGG 2100  
DB 1967 TCTCGAGACACTGTCCGACCGCTTTGGCGGCGCGCGAGTCCGTGCTTGGCTTGGCTACTTGG 2026  
QY 2101 AGCCACTATCGACTACCGCATATGCGGCGACCAACCGGCTCTGTGGATCTATCGAATCTA 2160  
DB 2027 AGCCACTATCGACTACCGCATATGCGGCGACCAACCGGCTCTGTGGATCTATCGAATCTA 2086  
QY 2161 AATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATPACGAACCTTAA 2220  
DB 2087 AATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATPACGAACCTTAA 2146  
QY 2221 CAGCANTGCGGTAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2280  
DB 2147 CAGCANTGCGGTAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2206  
QY 2281 TATGTTTCACTCCCTCAGGAGTTACGTCTGTGAAAGTATGAATCTTGGAAAGTTGCGAG 2340  
DB 2207 TATGTTTCACTCCCTCAGGAGTTACGTCTGTGAAAGTATGAATCTTGGAAAGTTGCGAG 2266  
QY 2341 TGTTAATCGGCTGATTAACCGGCAATCCGTACGTTGGCAAGTGTGTTGTTACCGG 2400  
DB 2267 TGTTAATCGGCTGATTAACCGGCAATCCGTACGTTGGCAAGTGTGTTGTTACCGG 2326  
QY 2401 AGGAGTAATCTCCAACTCTCTGGAGAGTAGGCACCAACAAACAGATCCAGCGTGT 2460  
DB 2327 AGGAGTAATCTCCAACTCTCTGGAGAGTAGGCACCAACAAACAGATCCAGCGTGT 2386  
QY 2461 GTACTTGATCAACATGAAGAAGCATTTCTGATTTGGAAGATCAAGTGTTCAGGAGGT 2520  
DB 2387 GTACTTGATCAACATGAAGAAGCATTTCTGATTTGCAAGATCAAGTGTTCAGGAGGT 2446  
QY 2521 ACTGATTGGACATTTCCAAAGCGTCTGCTAGGTTGCAACCGCATAGGTTGTAGAGTGTG 2580  
DB 2447 ACTGATTGGACATTTCCAAAGCGTCTGCTAGGTTGCAACCGCATAGGTTGTAGAGTGTG 2506  
QY 2581 CAATACACTTGGCTACAAATTTCAACCCCTTGGCAACTGCAAGTGTGTTGTTGAAACAGCAT 2640  
DB 2507 CAATACACTTGGCTACAAATTTCAACCCCTTGGCAACTGCAAGTGTGTTGTTGAAACAGCAT 2566  
QY 2641 CTTCAATCTGGCAAGCTCTGCTGCTATATCGACGCGCAACAGAAATCACTGGGAAT 2700  
DB 2567 CTTCAATCTGGCAAGCTCTGCTGCTATATCGACGCGCAACAGAAATCACTGGGAAT 2626  
QY 2701 CAATACACTTGGCTTGAAGAGAGTCTGAGGCAACGAAATCTGGATCAGCGTAT 2760  
DB 2627 CAATACACTTGGCTTGAAGAGAGTCTGAGGCAACGAAATCTGGATCAGCGTAT 2686  
QY 2761 TATCAGCAATTAACATTAAGAGAGGCGGCGGCGATGTCAATTAACAGGCGT 2820  
DB 2687 TATCAGCAATTAACATTAAGAGAGGCGGCGGCGATGTCAATTAACAGGCGT 2746  
QY 2821 ATGTGTCATTTGAAACCATCATCTTGGCAGCAGTAACGAACTGGTTTCTCGACCAAAATA 2880  
DB 2747 ATGTGTCATTTGAAACCATCATCTTGGCAGCAGTAACGAACTGGTTTCTCGACCAAAATA 2806

QY 2881 TTTTGTCACTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGCAGCTACTGCTGGGCGC 2940  
DB 2807 TTTTGTCACTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGCAGCTACTGCTGGGCGC 2866  
QY 2941 CTCTGTCTAGCAAGATACACTTAGCACCACCACTTGTGGGCAACGTAGATCACTTCTGGG 3000  
DB 2867 CTCTGTCTAGCAGATACACTTAGCA - CAACCTTGTGGGCAACGTAGATCACTTCTGGG 2925  
QY 3001 TAAGGCTACCATCTTCTTAGGTGGAGATGCAAAAACAAATTTCTTTGCAACAGCAACTT 3060  
DB 2926 TAAGGCTACCATCTTCTTAGGTGGAGATGCAAAAACAAATTTCTTTGCAACAGCAACTT 2985  
QY 3061 TGGCAGGAACACCCAGCATCAGGGAAGTGAAGGAGAAATTTGGGTTCCACAGGAATAT 3120  
DB 2986 TGGCAGGAACACCCAGCATCAGGGAAGTGAAGGAGAAATTTGGGTTCCACAGGAATAT 3045  
QY 3121 AGAGGCCAACTTTCTCAATAGTCTTGC AAAAGAGAGCAGACTACACAGGCGCAAGTCT 3180  
DB 3046 AGAGGCCAACTTTCTCAATAGTCTTGC AAAAGAGAGCAGACTACACAGGCGCAAGTCT 3105  
QY 3181 CAATTTGCAAAAGTCTCGTTAGTTAGCTTCAATGGAATTTCTTGAOCTTATCTATAGAGA 3240  
DB 3106 CAATTTGCAAAAGTCTCGTTAGTTAGCTTCAATGGAATTTCTTGAOCTTATCTATAGAGA 3165  
QY 3241 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGATAAGTTCCTCTGGGAAAGGAGCTT 3300  
DB 3166 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGATAAGTTCCTCTGGGAAAGGAGCTT 3225  
QY 3301 CTAACACAGGTGTCTTCAAAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGCGCT 3360  
DB 3226 CTAACACAGGTGTCTTCAAAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGCGCT 3285  
QY 3361 TGTCAACATTTGAGCAACATTTGCGAATTTGGTTGACTAATTCATATCTGTTCCG 3420  
DB 3286 TGTCAACATTTGAGCAACATTTGCGAATTTGGTTGACTAATTCATATCTGTTCCG 3345  
QY 3421 TTTTCTGATAGGACGACGAGGCGCATCTTCAATTTCTTGTGAGGAGGCTTGAAGACGT 3480  
DB 3346 TTTTCTGATAGGACGACGAGGCGCATCTTCAATTTCTTGTGAGGAGGCTTGAAGACGT 3405  
QY 3481 CAATTTTGCACAAATTCATACGACTTCAAGGGAATCTTCTTAGGTTTGGATTTCTTCTT 3540  
DB 3406 CAATTTTGCACAAATTCATACGACTTCAAGGGAATCTTCTTAGGTTTGGATTTCTTCTT 3465  
QY 3541 TAGGTTGTTCTTCTGTTGATCTGCTGGCTTGGCATCTCTTCTTCTTCTAGTACCTTTAGGG 3600  
DB 3466 TAGGTTGTTCTTCTGTTGATCTGCTGGCTTGGCATCTCTTCTTCTTCTAGTACCTTTAGGG 3525  
QY 3601 ACTTCATATCCAGGTTTCTTCCACCTCGTCCAAACGTCAACCGTACTTGGCAGCATCTAA 3660  
DB 3526 ACTTCATATCCAGGTTTCTTCCACCTCGTCCAAACGTCAACCGTACTTGGCAGCATCTAA 3585  
QY 3661 CTATGCAAAATAAAATAAGTACGACATTTCCAGGCTATATCTTCTTGGATTTAGCTT 3720  
DB 3586 CTATGCAAAATAAAATAAGTACGACATTTCCAGGCTATATCTTCTTGGATTTAGCTT 3645  
QY 3721 CTGCAAGTTCAATCAGCTTCTTCCCTTAATTTTAGGCTTCAACAAAACCTTCGTCTCAATA 3780  
DB 3646 CTGCAAGTTCAATCAGCTTCTTCCCTTAATTTTAGGCTTCAACAAAACCTTCGTCTCAATA 3705  
QY 3781 ACCGTTTGGTATAAGAACCTTCTGAGCATTTCTTACGATCCCAAGGTCGCTTCCA 3840  
DB 3706 ACCGTTTGGTATAAGAACCTTCTGAGCATTTCTTACGATCCCAAGGTCGCTTCCA 3765  
QY 3841 TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAACCAA 3900  
DB 3766 TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAACCAA 3825  
QY 3901 CACCTGTTTGTCAACCAAAATTTCAAGCAGTCTCCATCAATTCAAATTCGATACCCA 3960  
DB 3826 CACCTGTTTGTCAACCAAAATTTCAAGCAGTCTCCATCAATTCAAATTCGATACCCA 3885  
QY 3961 GCAACTTTTGTGCTCCAGATAGCACCTTTATACCAAAACCGGTGACGACGAGTT 4020

3886 GCAACCTTTTGGAGTCTCCAGATGAGCACTTTTATACCAAAACCGGTGACGAGATTT 3945  
Db  
4021 GTTAGACTCCAGTTTGTGTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4080  
QY  
3946 GGTAGACTCCAGTTTGTGTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4005  
Db  
4081 GSCCAACGAGTAATAGAGAGTCCAGCAACAAAGTAGTGAATAGACCATCGGGCGGT 4140  
QY  
4006 GSCCAACGAGTAATAGAGAGTCCAGCAACAAAGTAGTGAATAGACCATCGGGCGGT 4065  
Db  
4141 CAGTAGTCAAGACGCCCAACAAATTTTCACTGACAGGGAATTTTGTGACATCTTCAGAAA 4200  
QY  
4066 CAGTAGTCAAGACGCCCAACAAATTTTCACTGACAGGGAATTTTGTGACATCTTCAGAAA 4125  
Db  
4201 GTTCGTATTTCAGTAGTCAATTCGCGAGCATCAATTAATGGGATTTATACCAAGCAACAG 4260  
QY  
4126 GTTCGTATTTCAGTAGTCAATTCGCGAGCATCAATTAATGGGATTTATACCAAGCAACAG 4185  
Db  
4261 TGGAGTCAATCTACCAACTTTTGGGTTCTCAGAAAAGCATTAACAGTCTTACTACCGC 4320  
QY  
4186 TGGAGTCAATCTACCAACTTTTGGGTTCTCAGAAAAGCATTAACAGTCTTACTACCGC 4245  
Db  
4321 CATTAGTGAATCTTTCAATCGCCAGTGGAGAGAAAAGGCCACAGCGATCTAGCAT 4380  
QY  
4246 CATTAGTGAATCTTTCAATCGCCAGTGGAGAGAAAAGGCCACAGCGATCTAGCAT 4305  
Db  
4381 TAGGGGCAAGGATGCACTTTTATCAACAGGTCCTATAGATTAACCTAGCGCTGGGA 4440  
QY  
4306 TAGGGGCAAGGATGCACTTTTATCAACAGGTCCTATAGATTAACCTAGCGCTGGGA 4365  
Db  
4441 TCATCTTTGGCAACTCTTTCTGCAAACTTAGGTCCAAATCACTTCATTGTATACCAT 4500  
QY  
4366 TCATCTTTGGCAACTCTTTCTGCAAACTTAGGTCCAAATCACTTCATTGTATACCAT 4425  
Db  
4501 TATTGTACAACTTAGCAAGTTGTGATCAGCTCTCAAAATGGTCTCTCTTAACGGATG 4560  
QY  
4426 TATTGTACAACTTAGCAAGTTGTGATCAGCTCTCAAAATGGTCTCTCTTAACGGATG 4485  
Db  
4561 ACTCACTTGCACATTAATCTGAGCTCAGTGAATGAGTGAATCTGATGAGTTGTGCA 4620  
QY  
4486 ACTCACTTGCACATTAATCTGAGCTCAGTGAATGAGTGAATCTGATGAGTTGTGCA 4545  
Db  
4621 GCTGTGTCAGCAGATAGGAAACACGGCTTTTCTACCAAACTCAAGGAATTTACAACT 4680  
QY  
4546 GCTGTGTCAGCAGATAGGAAACACGGCTTTTCTACCAAACTCAAGGAATTTACAACT 4605  
Db  
4681 CTGCAACACTTGGGTATGAGTAGCAAGGAAATGTATATCTGAGTTCGAGTCAGTGA 4740  
QY  
4606 CTGCAACACTTGGGTATGAGTAGCAAGGAAATGTATATCTGAGTTCGAGTCAGTGA 4665  
Db  
4741 TGTAGTCTTGAGAAATCTGAAAGCGTATTTTATATCACTGAGTTCAGTCACTCAGGAGA 4800  
QY  
4666 TGTAGTCTTGAGAAATCTGAAAGCGTATTTTATATCACTGAGTTCAGTCACTCAGGAGA 4725  
Db  
4801 TCTCTACCGCGGACGATCTGTGGCGGA----- 4828  
QY  
4726 TCTCTACCGCGGACGATCTGTGGCGGACCTTCAGGGGGGGGGGGGGGGGGGGGGGGGG 4785  
Db  
4829 ----- 4828  
QY  
4786 CTCGTGAAGAGGTGTGTGTAATCATACCGGCTGATTCGCCCATCATCCAGCCAGA 4845  
Db  
4829 ----- 4828  
QY  
4846 AAGTAGGGGACCAAGTTGATGAGAGCTTTGTTGTTAGTGGACCAAGTTGATTTTGA 4905  
Db  
4829 ----- 4828  
QY  
4906 ACTTTTGTCTTGGCCAGGAACGGTCTGCGTGTCTGGGAAGATGCGTGATCTGATCCCTCA 4965  
Db  
4829 ----- 4828  
QY

4966 ACTCAGCAAAAGTTTCGATTTTATTCACAAAGCCGCGTCCCGTCAAGTCAGCGTAATGCT 5025  
Db  
4829 ----- 4828  
QY  
5026 CTGCCAGTGTACAAACCAATTAACCAATTTCTGATTAGAAAACTCATCGAGCATCAAAATG 5085  
Db  
4829 ----- 4828  
QY  
5086 AAACCTGCAATTTTATTCATATCAGGATTATCAATACCATATTTTGA AAAAGCGGTTTCG 5145  
Db  
4829 ----- 4828  
QY  
5146 TAATGAAGGAGAAAACCTCACCGAGGAGTTCCATAGTAGGCAAGATCCTGATCGGTC 5205  
Db  
4829 ----- 4828  
QY  
5206 TGGGATTCGAGCTGCTCCAAATCAATACCACTATTAATTTCCCTCTGTC AAAAATAG 5265  
Db  
4829 ----- 4828  
QY  
5266 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGCGAAAAGCTT 5325  
Db  
4829 ----- 4828  
QY  
5326 ATGCATTTCTTTCCAGACTTTGTTCAACAGGCCAGCCATTAACGCTCGTCAATCAAAATCACT 5385  
Db  
4829 ----- 4828  
QY  
5386 CGCATCAACCAAAACCGTTTATTCATTCGTGATTGCGCCTGAGCGAGACGAATACGCGATC 5445  
Db  
4829 ----- 4828  
QY  
5446 GCTGTTAAAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCGCAGGAACACTGCCAG 5505  
Db  
4829 ----- 4828  
QY  
5506 CGCATCAACAAATTTTCACTGATCAGGATATTCCTTAATACCTGGAATGCTGTGTTT 5565  
Db  
4829 ----- 4828  
QY  
5566 CCGCGGATCGCAGTGGTGAGTAAACCATCATCAGGAGTACGGAATAAAATGCTTGAT 5625  
Db  
4829 ----- 4828  
QY  
5626 GGTGGAAGAGGATAAATTCGCTGAGCCAGTTTAGTCTGACCATCTCATCTGTAACATC 5685  
Db  
4829 ----- 4828  
QY  
5686 ATTGGCAACGCTACCTTTGCCATGTTTTCAGAAAACAACTCTGCGCATCGGGCTTCCCAT 5745  
Db  
4829 ----- 4828  
QY  
5746 CAATCGATAGATTGTCGACCTGANTGCCGACATTAATCGCGAGCCATTATACCCATA 5805  
Db  
4829 ----- 4828  
QY  
5806 TAAATCAGCATCCATGTTGGAATTTAATCGCGCCTCGAGCAAGAGGTTTCCCGTTGAAT 5865  
Db  
4829 ----- 4828  
QY  
5866 ATGGCTCATAAACCCCTTGTATTAATGTTTATGTAAGCAGACAGTTTATTTTCATGA 5925  
Db  
4829 ----- 4828  
QY  
5926 TGATATATTTTATCTTGTGCAATGTAAATCAGAGATTTTGGACACAACTGGCTTTC 5985  
Db  
4829 ----- CCTCAGGTCCGCAATCACCGCGCCACAGGTGCGGTTCTGCGGCTTATAT 4879  
QY  
5986 CCCCCCCCCCTCAGGTCCGCAATCACCGCGCCACAGGTGCGGTTCTGCGGCTTATAT 6045  
Db  
4880 CCGCGACATCACCGATGGGAGATCGGCTCGCACTTCGGCTCATGAGCGCTGTTT 4939  
QY  
6046 CCGCGACATCACCGATGGGAGATCGGCTCGCACTTCGGCTCATGAGCGCTGTTT 6105  
Db

Qy 4940 CGCGTGGGTATGTTGGCAGGCCCCGTCGCGGGGACATGTTGGGGGCCCATCTCCTTGA 4999  
Db 6106 CGCGTGGGTATGTTGGCAGGCCCCGTCGCGGGGACATGTTGGGGGCCCATCTCCTTGA 6165  
Qy 5000 TGCACCAATTCCTTGGCGCGCGGTGCTCAACGCCCTCAACCTACTACTGGGCTGCTTCT 5059  
Db 6166 TGCACCAATTCCTTGGCGCGCGGTGCTCAACGCCCTCAACCTACTACTGGGCTGCTTCT 6225  
Qy 5060 AATGACGAGTTCGATTAAGGAGGAGGCTGAGTATCTATGATTTGGAAGTATGGGAATGGT 5119  
Db 6226 AATGACGAGTTCGATTAAGGAGGAGGCTGAGTATCTATGATTTGGAAGTATGGGAATGGT 6285  
Qy 5120 GATACCCGCAATTCCTGAGTCTTTGAGTCTCTATCAGATATATGCCCACTAAAGCAAC 5179  
Db 6286 GATACCCGCAATTCCTGAGTCTTTGAGTCTCTATCAGATATATGCCCACTAAAGCAAC 6345  
Qy 5180 CGAGGAGGAGATTCATGTTAAATTCCTGACTTTTGGTCTATCAGTAGACTCGAACTG 5239  
Db 6346 CGAGGAGGAGATTCATGTTAAATTCCTGACTTTTGGTCTATCAGTAGACTCGAACTG 6405  
Qy 5240 TGAGACTATCTCGGTTATGACACAGAAATGCTCTTCTTGGAGACAGTAATGAAGTCCC 5299  
Db 6406 TGAGACTATCTCGGTTATGACACAGAAATGCTCTTCTTGGAGACAGTAATGAAGTCCC 6465  
Qy 5300 ACCAATAAGAAATCCTTGTATCAGGAACAAACTCTTGTCTCGAACTTTTTCGGTGCC 5359  
Db 6466 ACCAATAAGAAATCCTTGTATCAGGAACAAACTCTTGTCTCGAACTTTTTCGGTGCC 6525  
Qy 5360 TTGAACTATAAATGATGATGATGATGCGGTAGGAATGAGCGGGCAATGCTTACC 5419  
Db 6526 TTGAACTATAAATGATGATGATGCGGTAGGAATGAGCGGGCAATGCTTACC 6585  
Qy 5420 TTCTGGACCTTCAAGAGGATGATGAGGTTTGTAGACTGATGCGCAACTTTCAGTGACAA 5479  
Db 6586 TTCTGGACCTTCAAGAGGATGATGAGGTTTGTAGACTGATGCGCAACTTTCAGTGACAA 6645  
Qy 5480 GTTGCTATTCGTTCAACCAATTCGAAATCCAGAGAAATCAAGTTGTTGTCFATATT 5539  
Db 6646 GTTGCTATTCGTTCAACCAATTCGAAATCCAGAGAAATCAAGTTGTTGTCFATATT 6705  
Qy 5540 GATCCAGCCAGTGGGCTTGTAAACTGACAACTAGTGCTGCTGTTTGGGTGCTATT 5599  
Db 6706 GATCCAGCCAGTGGGCTTGTAAACTGACAACTAGTGCTGCTGTTTGGGTGCTATT 6765  
Qy 5600 TGATGAAATTAATCTAGTCTTGTATCTAAATTAATCTTGAAGCAAGTATGCTGCTGTTTGAAGTCAATACC 5659  
Db 6766 TGATGAAATTAATCTAGTCTTGTATCTAAATTAATCTTGAAGCAAGTATGCTGCTGTTTGAAGTCAATACC 6825  
Qy 5660 CAAATCTAAACTCTTTTAAACGTTTAAAGGACAAAGTATGCTGCTGTTTAAACCCC 5719  
Db 6826 CAAATCTAAACTCTTTTAAACGTTTAAAGGACAAAGTATGCTGCTGTTTAAACCCC 6885  
Qy 5720 AAATCAGCTCGTAGTCTGATCTCATCAACTTGAAGGCACTATCTTGTGTTTGAAGAAAT 5779  
Db 6886 AAATCAGCTCGTAGTCTGATCTCATCAACTTGAAGGCACTATCTTGTGTTTGAAGAAAT 6945  
Qy 5780 TTGGGAGATGCGATATCGAGAAAAAGTACGCTGATTTTAAACGTTGAATTTATCTCAA 5839  
Db 6946 TTGGGAGATGCGATATCGAGAAAAAGTACGCTGATTTTAAACGTTGAATTTATCTCAA 7005  
Qy 5840 GATCTGCTCGCGCTTTCGTTGATGACGTTGAAAACTCTGACAACTGAGCTCCCG 5899  
Db 7006 GATCTGCTCGCGCTTTCGTTGATGACGTTGAAAACTCTGACAACTGAGCTCCCG 7065  
Qy 5900 GAGACGCTCAGCTTCTGTTAAGGATGCGGAGCAGACAAAGCCGTCAGGCGCG 5959  
Db 7066 GAGACGCTCAGCTTCTGTTAAGGATGCGGAGCAGACAAAGCCGTCAGGCGCG 7125  
Qy 5860 TCAGCGGCTGTTGGCGGCTTCGGGCGGAGCATGACCCAGTCAAGTACGATACGGA 6019  
Db 7126 TCAGCGGCTGTTGGCGGCTTCGGGCGGAGCATGACCCAGTCAAGTACGATACGGA 7185

Qy 6020 GTGTATATGCTTAACTATGCGGCATCAGACAGATTGTTACTGAGATGACCAATATGC 6079  
Db 7186 GTGTATATGCTTAACTATGCGGCATCAGACAGATTGTTACTGAGATGACCAATATGC 7245  
Qy 6080 GGTGTGAATACCGACACAGATCGGTAAGAGAAAAATACCGATCAGGGCTCTTCGGCTT 6139  
Db 7246 GGTGTGAATACCGACACAGATCGGTAAGAGAAAAATACCGATCAGGGCTCTTCGGCTT 7305  
Qy 6140 CTTCTGCTACTGATCGCTGCTCGGCTCGGCTCGGCGAGCGGTATCAGCTCACT 6199  
Db 7306 CTTCTGCTACTGATCGCTGCTCGGCTCGGCTCGGCGAGCGGTATCAGCTCACT 7365  
Qy 6200 CAAAGCGGTATACCGTTATCCACAGATCAGGGATTAACGACAGAAAGACATGTGAG 6259  
Db 7366 CAAAGCGGTATACCGTTATCCACAGATCAGGGATTAACGACAGAAAGACATGTGAG 7425  
Qy 6260 CAAAAGCGCAGCAAAAAGGCCAGAACCGTTAAAGAGCGCGCTTGTGCGCTTTTCCATA 6319  
Db 7426 CAAAAGCGCAGCAAAAAGGCCAGAACCGTTAAAGAGCGCGCTTGTGCGCTTTTCCATA 7485  
Qy 6320 GGTCTCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACC 6379  
Db 7486 GGTCTCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACC 7545  
Qy 6380 CGACAGACTATAAGATACCGAGGCTTTCCCGCTCGAAGCTCCCTCGTGGCTCTCCTG 6439  
Db 7546 CGACAGACTATAAGATACCGAGGCTTTCCCGCTCGAAGCTCCCTCGTGGCTCTCCTG 7605  
Qy 6440 TTCCGACCTCGGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAGCGTGGCG 6499  
Db 7606 TTCCGACCTCGGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAGCGTGGCG 7665  
Qy 6500 TTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGCTGTAGTCTGTTGCTTCCAAAGCTGG 6559  
Db 7666 TTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGCTGTAGTCTGTTGCTTCCAAAGCTGG 7725  
Qy 6560 GCTGTGTGACGAAACCCCGTTACCGCGGACCGCTGCGCTTATCCGTTAATCTCCTC 6619  
Db 7726 GCTGTGTGACGAAACCCCGTTACCGCGGACCGCTGCGCTTATCCGTTAATCTCCTC 7785  
Qy 6620 TTGAGTCCAAACCCCGTTAAGACACGCTTATCCCACTTGGCAGCAGCCACTGTTAAACAGGA 6679  
Db 7786 TTGAGTCCAAACCCCGTTAAGACACGCTTATCCCACTTGGCAGCAGCCACTGTTAAACAGGA 7845  
Qy 6680 TTGACGAGCGGATGATGAGCGGTGCTACAGATTCTTGAAGTGTGGCTTAACTACG 6739  
Db 7846 TTGACGAGCGGATGATGAGCGGTGCTACAGATTCTTGAAGTGTGGCTTAACTACG 7905  
Qy 6740 GCTACACTAAGAGGACAGTATTTGTTATCTGCGCTCTCTGAAAGCCAGTTTACCTTCGGAA 6799  
Db 7906 GCTACACTAAGAGGACAGTATTTGTTATCTGCGCTCTCTGAAAGCCAGTTTACCTTCGGAA 7965  
Qy 6800 AAAGAGTGTGATCTTGTATCGGCAACAAACCCAGCTGGTAGCGGTGTTTTTTG 6859  
Db 7966 AAAGAGTGTGATCTTGTATCGGCAACAAACCCAGCTGGTAGCGGTGTTTTTTG 8025  
Qy 6860 TTTGCAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAGATCTTGTGATCTTTT 6919  
Db 8026 TTTGCAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAGATCTTGTGATCTTTT 8085  
Qy 6920 CTAAGGGTCTGACGCTCAGTGAAGAAACTCAGTTTGAAGGATTTTGTGATGAGAT 6979  
Db 8086 CTAAGGGTCTGACGCTCAGTGAAGAAACTCAGTTTGAAGGATTTTGTGATGAGAT 8145  
Qy 6980 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAATCT 7039  
Db 8146 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAATCT 8205  
Qy 7040 AAAGTATATAGGTAACCTTGTCTGACGTTACATGCTTAAATGCTTAACTAGTGGGACCTA 7099  
Db 8206 AAAGTATATAGGTAACCTTGTCTGACGTTACATGCTTAAATGCTTAACTAGTGGGACCTA 8265  
Qy 7100 TCTCAGGATCTGCTTATTTGCTTCACTCATAGTTGCGCTGACTCCCGCTCGTAGATAA 7159





QY 1350 ----- 1349  
Db 1679 CAATTCGACAACTCTCAAAACCGAAGTTGCGCTACCTTGGCATACCGCTTCGGCAACGT 1738  
QY 1350 ----- 1349  
Db 1739 AACGCCCGAGTTTCCTTAAGCCCAACGCTTCAAGGTTTGGTTGATGATGACAGACATAGG 1798  
QY 1350 -----CGGATTCGCCGT 1360  
Db 1799 CAACGAATACGACCAAGTGGTTGTCGGCGGAATACGACTTCTCCAAACGCACTTCTGC 1858  
QY 1361 GGTGCTTCCACCCCATCGGCATCGAAGCACTCAAGAAAGAAATGTCCTTCTAATAG 1420  
Db 1859 CTTGGTTTCTGCCGTTGTTGCAAGAGGCAAGGCGAATAAATTCGTACGCACTGC 1918  
QY 1421 TGGCGTAGAATTCCTAGGGCGCGCAATTAATTCGCTTAGACATGACTGTTCTCA 1480  
Db 1919 CGCGGTGTGCTGCGCCCAAAATCTAATCTGGATCCTTAGACATGACTGTTCTCA 1978  
QY 1481 GTTCAAGTTGGCACTTACGAGAAAGCCGTTCTGCTAGATTTCTAATCAAGAGGATGTCA 1540  
Db 1979 GTTCAAGTTGGCA-TTACGAGAAAGCCGCTTCTGCTAGATTTCTAATCAAGAGGATGTCA 2037  
QY 1541 GAATGCCATTTGCGTGAGAGTACAGCTTCATTTTGTATCTTTTATTTTGTAACTTA 1600  
Db 2038 GAATGCCATTTGCGTGAGAGTACAGCTTCATTTTGTATCTTTTATTTTGTAACTTA 2097  
QY 1601 TATAGTATAGGATTTTCTCATTTGTTTCTCTCTGACGAGTGTCTCTGATCAG 1660  
Db 2098 TATAGTATAGGATTTTCTCATTTGTTTCTCTCTGACGAGTGTCTCTGATCAG 2157  
QY 1661 CCTATCTCGAGCTGATGAATATCTTGTGTAGGGTTTGGGAAATCAATTCAGATTGA 1720  
Db 2158 CCTATCTCGAGCTGATGAATATCTTGTGTAGGGTTTGGGAAATCAATTCAGATTGA 2217  
QY 1721 TGTCTTCTCGGATTTTCTCCACTCTCTCAGAGTACAGAAATTAAGTCAGAAATTCGT 1780  
Db 2218 TGTCTTCTCGGATTTTCTCCACTCTCTCAGAGTACAGAAATTAAGTCAGAAATTCGT 2277  
QY 1781 TTGTCAAGCTTATCGATTAAGCTTTAATGCGGTAGTTTATCAAGTTAAATTTGCTAACGC 1840  
Db 2278 TTGTCAAGCTTATCGATTAAGCTTTAATGCGGTAGTTTATCAAGTTAAATTTGCTAACGC 2337  
QY 1841 AGTCAGGACCGTGTATGAATCTAAGAAATGCGCTCATCTGCTCGGACCGTCAAC 1900  
Db 2338 AGTCAGGACCGTGTATGAATCTAAGAAATGCGCTCATCTGCTCGGACCGTCAAC 2397  
QY 1901 CTGGATGCTGTAGGATAGGCTTGGTTATGCGGTACTGCGGCGCTCTTTCGGGATATC 1960  
Db 2398 CTGGATGCTGTAGGATAGGCTTGGTTATGCGGTACTGCGGCGCTCTTTCGGGATATC 2457  
QY 1961 GTCCATTCGACAGATCGCCAGTCACTATGCGGTGCTGTAGCGCTATATGCTGTATG 2020  
Db 2458 GTCCATTCGACAGATCGCCAGTCACTATGCGGTGCTGTAGCGCTATATGCTGTATG 2517  
QY 2021 CAATTTCTATGCGACCGCTTCTCGGAGCACTGTCGACCGCTTTGGCCGCGCCCAATC 2080  
Db 2518 CAATTTCTATGCGACCGCTTCTCGGAGCACTGTCGACCGCTTTGGCCGCGCCCAATC 2577  
QY 2081 CTGCTCGCTTCTGCTACTTGGAGCACTATCGACTACCGGATCTGCGGACCAACCGCTC 2140  
Db 2578 CTGCTCGCTTCTGCTACTTGGAGCACTATCGACTACCGGATCTGCGGACCAACCGCTC 2637  
QY 2141 CTGTGGATCTATCGAATCTAATGTAGTTAAATCTCTAATAATTAATTAAGTCCCGAG 2200  
Db 2638 CTGTGGATCTATCGAATCTAATGTAGTTAAATCTCTAATAATTAATTAAGTCCCGAG 2697  
QY 2201 TTTCTCCATAGCACTTAAAGCAATGCGGTGAGCATCTAGACCTTCAACAGCAGCCAG 2260  
Db 2698 TTTCTCCATAGCACTTAAAGCAATGCGGTGAGCATCTAGACCTTCAACAGCAGCCAG 2757

QY 2261 ATCCATCTACTGCTTGGCCAAATATGTTTCAGTCCCTCAGAGTTACGCTTGTGAGGTGAT 2320  
Db 2758 ATCCATCTACTGCTTGGCCAAATATGTTTCAGTCCCTCAGAGTTACGCTTGTGAGGTGAT 2817  
QY 2321 GAACTTCTGGAAGGTTGAGTGTAACTCCGCTGATTTAGCGGCATATCGTACGTTGG 2380  
Db 2818 GAACTTCTGGAAGGTTGAGTGTAACTCCGCTGATTTAGCGGCATATCGTACGTTGG 2877  
QY 2381 AAAAGTGTGTTGATCCGAGGAGTATCTCCAACTCTCTCGAGAGTAGGCAACCAAC 2440  
Db 2878 AAAAGTGTGTTGATCCGAGGAGTATCTCCAACTCTCTCGAGAGTAGGCAACCAAC 2937  
QY 2441 AAACACAGATCCAGGCTGTGTACTTGTATCAATCAATCAATCAATCAATCAATCAAT 2500  
Db 2938 AAACACAGATCCAGGCTGTGTACTTGTATCAATCAATCAATCAATCAATCAATCAAT 2997  
QY 2501 GATCAAGTGTTCAGAGGCTACTGATTTGGAATTTCCAAAGCCTGCTGTAAGTTGCAAC 2560  
Db 2998 GATCAAGTGTTCAGAGGCTACTGATTTGGAATTTCCAAAGCCTGCTGTAAGTTGCAAC 3057  
QY 2561 CGATAGGTTGTAGAGTGTGAATACACTTGGTACAAATTTCAACCTTGGCAACTGCAC 2620  
Db 3058 CGATAGGTTGTAGAGTGTGAATACACTTGGTACAAATTTCAACCTTGGCAACTGCAC 3117  
QY 2621 AGCTTGGTTGTGAACAGCATCTTCAATTTCTGGCAAGCTCTTGTCTGTCTATCGACAGC 2680  
Db 3118 AGCTTGGTTGTGAACAGCATCTTCAATTTCTGGCAAGCTCTTGTCTGTCTATCGACAGC 3177  
QY 2681 AAACAGATCACTTGGGAATCAATACCAATGTTACGTTGACAGAGAGGTTCTGAGGCAAC 2740  
Db 3178 AAACAGATCACTTGGGAATCAATACCAATGTTACGTTGACAGAGAGGTTCTGAGGCAAC 3237  
QY 2741 GAAATCTGGATCAGGCTATTTATCAGCAATTAATAGAACTTTCAGAAAGCCCGAGGAGCAT 2800  
Db 3238 GAAATCTGGATCAGGCTATTTATCAGCAATTAATAGAACTTTCAGAAAGCCCGAGGAGCAT 3297  
QY 2801 GTCAATATACACAGGCTGATGTCTATTTTGAACCAATCATCTTGGCAGCAGTAACGAA 2860  
Db 3298 GTCAATATACACAGGCTGATGTCTATTTTGAACCAATCATCTTGGCAGCAGTAACGAA 3357  
QY 2861 CTGGTTTCTCGAACCAATATTTTGTACACTTAGGAACAGTTTCTGTTCCGTAAAGCAT 2920  
Db 3358 CTGGTTTCTCGAACCAATATTTTGTACACTTAGGAACAGTTTCTGTTCCGTAAAGCAT 3417  
QY 2921 AGCAGCTACTGCTCGGCGCTCTCTGTAGCAAGATACACTTAGCAACCACTTGTGGGC 2980  
Db 3418 AGCAGCTACTGCTCGGCGCTCTCTGTAGCAAGATACACTTAGCAACCACTTGTGGGC 3477  
QY 2981 AACGTAGATGACTTCTGGGTAAAGGTACCACTCTTCTTAGGTGAGATGCAAAACAAAT 3040  
Db 3478 AACGTAGATGACTTCTGGGTAAAGGTACCACTCTTCTTAGGTGAGATGCAAAACAAAT 3537  
QY 3041 TTCTTTGCAACCCAGCAAACTTTTGGCAGGAACCCAGCATCAGGGAAGTGGGAAGCAGAAAT 3100  
Db 3538 TTCTTTGCAACCCAGCAAACTTTTGGCAGGAACCCAGCATCAGGGAAGTGGGAAGCAGAAAT 3597  
QY 3101 TCGGTTTCCACAGGAATATAGAGCCCACTTCTCAATAGGTCTTGCAGAAACCGAGACA 3160  
Db 3598 TCGGTTTCCACAGGAATATAGAGCCCACTTCTCAATAGGTCTTGCAGAAACCGAGACA 3657  
QY 3161 GACTACACAGGGAAGTCTCAACTTGTCAAGCTCTCCGTTAGTTGAGCTTTCATGAAATTT 3220  
Db 3658 GACTACACAGGGAAGTCTCAACTTGTCAAGCTCTCCGTTAGTTGAGCTTTCATGAAATTT 3717  
QY 3221 CTGAGCTTATCTATAGAGATCAATGGCTCTCTTAACGTTATCTGCAATTTGCAATAG 3280  
Db 3718 CTGAGCTTATCTATAGAGATCAATGGCTCTCTTAACGTTATCTGCAATTTGCAATAG 3777  
QY 3281 TTCTCTTGGGAAGGAGCTTCTTAACAGAGGTGTCTTCAAGCGACTCCCAACTTTGGC 3340  
Db 3778 TTCTCTTGGGAAGGAGCTTCTTAACAGAGGTGTCTTCAAGCGACTCCCAACTTTGGC 3837  
QY 3341 AGTATGTTCTAAAGGGCTTTGTACCACTTTTGTAGCAACATTTGTGCAAAATTTGTTGAC 3400



Db	3838	AGTTAGTTCTAAAGGGCTTTGTACCACTTTGACGAACATTTGCGAATTTGTTTGAC	3897
Qy	3401	TAATTCCTAATCTGTTCCGTTTCTGGATAGGACGACGAGGCACTTCAATTTCTTG	3460
Db	3898	TAATTCCTAATCTGTTCCGTTTCTGGATAGGACGACGAGGCACTTCAATTTCTTG	3957
Qy	3461	TGAGGAGCCTTAGAAACGTCATTTTGTCACAAATTAATAGACCTTCAGAAAGGACTTC	3520
Db	3958	TGAGGAGCCTTAGAAACGTCATTTTGTCACAAATTAATAGACCTTCAGAAAGGACTTC	4017
Qy	3521	TTTAGGTTTGGATCTTCTTTAGGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTT	3580
Db	4018	TTTAGGTTTGGATCTTCTTTAGGTTTGTCTGCTGCTGCTGCTGCTGCTGCTT	4077
Qy	3581	CCTTCTAGTGACCTTTAGGGACTTCATATCAGAGTTTCTCTCCACCTCGTCCAAAGTCAC	3640
Db	4078	CCTTCTAGTGACCTTTAGGGACTTCATATCAGAGTTTCTCTCCACCTCGTCCAAAGTCAC	4137
Qy	3641	ACCGTACTTGGACATCTAATCTAATGCAAAATAAATAAGTCAGACATTTCCCAAGGCTAT	3700
Db	4138	ACCGTACTTGGACATCTAATCTAATGCAAAATAAATAAGTCAGACATTTCCCAAGGCTAT	4197
Qy	3701	ATCTTCTTGGATTTAGTCTTCTGCAAGTTCTATCAGCTTCTCTCCCTTAATTTTAGGCTTCAA	3760
Db	4198	ATCTTCTTGGATTTAGTCTTCTGCAAGTTCTATCAGCTTCTCTCCCTTAATTTTAGGCTTCAA	4257
Qy	3761	CAAACTTCGCTGCTCAAAATAACCGTTTGGTATAAGAACCTTTCTGGAGCATTTGCTCTTACG	3820
Db	4258	CAAACTTCGCTGCTCAAAATAACCGTTTGGTATAAGAACCTTTCTGGAGCATTTGCTCTTACG	4317
Qy	3821	ATCCCAAGGTGGCTTCCATGGCTCTAAGACCTTTGATTTGGCCCAAAACAGGAGTGGC	3880
Db	4318	ATCCCAAGGTGGCTTCCATGGCTCTAAGACCTTTGATTTGGCCCAAAACAGGAGTGGC	4377
Qy	3881	TTCCAAGTGACAGAAACCAACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATC	3940
Db	4378	TTCCAAGTGACAGAAACCAACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATC	4437
Qy	3941	ACAATCCAATTCGATACCCAGCACTTTTGTAGTTGCTCCAGATGACACCTTTATACCA	4000
Db	4438	ACAATCCAATTCGATACCCAGCACTTTTGTAGTTGCTCCAGATGACACCTTTATACCA	4497
Qy	4001	CAAAACGCTGACGACAGATTTGCTAGACTCCAGTTTGTCTCTTATAGCTCCGGAATAGA	4060
Db	4498	CAAAACGCTGACGACAGATTTGCTAGACTCCAGTTTGTCTCTTATAGCTCCGGAATAGA	4557
Qy	4061	CTTTTGGACGAGTACACAGCCCAACGAGTAATTAAGAGTACGACCAACCAAGTAGT	4120
Db	4558	CTTTTGGACGAGTACACAGCCCAACGAGTAATTAAGAGTACGACCAACCAAGTAGT	4617
Qy	4121	GAATAGACCATCGGGCGGTGAGTACAGTCAAGAGCCCAACAAATTTTCACTGACAGGAA	4180
Db	4618	GAATAGACCATCGGGCGGTGAGTACAGTCAAGAGCCCAACAAATTTTCACTGACAGGAA	4677
Qy	4181	CTTTTGGACATCTTCAGAAAGTTCGATTTAGTACGATCAATTTCCGAGCATCAATATGGG	4240
Db	4678	CTTTTGGACATCTTCAGAAAGTTCGATTTAGTACGATCAATTTCCGAGCATCAATATGGG	4737
Qy	4241	GATTTATACAGAAAGCAACAGTGGAGTACATCTACCACTTTTGGGTTCTCAGAAAGC	4300
Db	4738	GATTTATACAGAAAGCAACAGTGGAGTACATCTACCACTTTTGGGTTCTCAGAAAGC	4797
Qy	4301	ATAAAACAGTTCTPACTACCGCCATTAGTGAACCTTTTCAAAATCGCCAGTGAGAGAA	4360
Db	4798	ATAAAACAGTTCTPACTACCGCCATTAGTGAACCTTTTCAAAATCGCCAGTGAGAGAA	4857
Qy	4361	AGGCAAGGATPACTAGCATTTAGCGGGCAAGATGCAACTTTTATCAACAGGGTCTTATA	4420
Db	4858	AGGCAAGGATPACTAGCATTTAGCGGGCAAGATGCAACTTTTATCAACAGGGTCTTATA	4917
Qy	4421	GATTAACCTTAGCGCTGGATCATCTTTTGACCACTCTTTCTGCGCAATCTAGTGCCAA	4480

Db	4918	GATAACCCCTAGCCCTGGGATCATCTTTTGGCAAACTCTTTCTGCAAAATCTAGGTCAA	4977
Qy	4481	AATCACTTCAATGATACCACTTATTTGACAACTTTGACAAAGTTTGTGATCAGTCTCTCAA	4540
Db	4978	AATCACTTCAATGATACCACTTATTTGACAACTTTGACAAAGTTTGTGATCAGTCTCTCAA	5037
Qy	4541	TTGGTCTCTGTAAACGATGATCTCAACTTTGACATTAATTAAGCTCAGTCTGATGAT	4600
Db	5038	TTGGTCTCTGTAAACGATGATCTCAACTTTGACATTAATTAAGCTCAGTCTGATGAT	5097
Qy	4601	GAACTTGATCAGTTTGTGAGCTGGTTCAGCAGATAGGAAACACGGCTTTTCTTACCAA	4660
Db	5098	GAACTTGATCAGTTTGTGAGCTGGTTCAGCAGATAGGAAACACGGCTTTTCTTACCAA	5157
Qy	4661	ACTCAAGGAATTAATCAAACTCTGCAACCTTGGTATGAGTAGCAAGGAAATGTCAT	4720
Db	5158	ACTCAAGGAATTAATCAAACTCTGCAACCTTGGTATGAGTAGCAAGGAAATGTCAT	5217
Qy	4721	ACTTGAAGTCGGACAGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTATATCA	4780
Db	5218	ACTTGAAGTCGGACAGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTATATCA	5277
Qy	4781	GTGAGTCAGTCATCAGGAGATCTCTACGCCCGAGCATCTGTCGGCGACCTCGAGTCGG	4840
Db	5278	GTGAGTCAGTCATCAGGAGATCTCTACGCCCGAGCATCTGTCGGCGACCTCGAGTCGG	5325
Qy	4841	CATCACCGGGCCACAGGTGCGGTTGCTGCGGCTTATATCGCGACATCACCGATGGGA	4900
Db	5326	CATCACCGGGCCACAGGTGCGGTTGCTGCGGCTTATATCGCGACATCACCGATGGGA	5385
Qy	4901	AGATCGGGCTCGCACCTTCGGGCTCATGAGCGCTTTTTCGGCGTGGGTATGTTGTCAGG	4960
Db	5386	AGATCGGGCTCGCACCTTCGGGCTCATGAGCGCTTTTTCGGCGTGGGTATGTTGTCAGG	5445
Qy	4961	CCCCGTGGCGGGGACTGTTGGGCGCCATCTCTCTTGCATGACACCATCTCTTGGCGGGC	5020
Db	5446	CCCCGTGGCGGGGACTGTTGGGCGCCATCTCTCTTGCATGACACCATCTCTTGGCGGGC	5505
Qy	5021	GGTGTCAACGGCTCAACCTACTACTGGGCTGCTTCTTAATGACGAGTTCGATTAAGG	5080
Db	5506	GGTGTCAACGGCTCAACCTACTACTGGGCTGCTTCTTAATGACGAGTTCGATTAAGG	5565
Qy	5081	AGAGCTCGAGTATCTATGATTGGAAGTATGGAATGTTGATACCCGCAATCTTTCAGTGT	5140
Db	5566	AGAGCTCGAGTATCTATGATTGGAAGTATGGAATGTTGATACCCGCAATCTTTCAGTGT	5625
Qy	5141	CTTGAGTCTCTTATCAGATTATGCCAACTAAAGCAACCGGAGGAGATTTTATGAT	5200
Db	5626	CTTGAGTCTCTTATCAGATTATGCCAACTAAAGCAACCGGAGGAGATTTTATGAT	5685
Qy	5201	AAATTTCTGACTTTTGTGTCATCAGTACCTCGAATCTGAGACTATCTCGGTTATGAC	5260
Db	5686	AAATTTCTGACTTTTGTGTCATCAGTACCTCGAATCTGAGACTATCTCGGTTATGAC	5745
Qy	5261	AGCAGAAATGTCCTTCTTGAGACATTAATGAAGTCCACCAATTAAGAAATCTTGT	5320
Db	5746	AGCAGAAATGTCCTTCTTGAGACATTAATGAAGTCCACCAATTAAGAAATCTTGT	5805
Qy	5321	ATCAGAAACAACTCTTCTTGTGTCATTTTTCGGTCTTGAATCTATAAAATGATAGT	5380
Db	5806	ATCAGAAACAACTCTTCTTGTGTCATTTTTCGGTCTTGAATCTATAAAATGATAGT	5865
Qy	5381	GGATATGTCGGTGAAGATGAGCGGGCAAAATGCTTACCTTCTGACCTTCAAGAGGTAT	5440
Db	5866	GGATATGTCGGTGAAGATGAGCGGGCAAAATGCTTACCTTCTGACCTTCAAGAGGTAT	5925
Qy	5441	GTAGGTTTGTAGATCTGATCCACTTCAGTGACAACTGCTATTTTCGTTCAACCA	5500
Db	5926	GTAGGTTTGTAGATCTGATCCACTTCAGTGACAACTGCTATTTTCGTTCAACCA	5985
Qy	5501	TTCCGAATCCAGAAATCAAAGTTGTTTGTCTACTATTTGATCCAAAGCAGTGGGTCTT	5560
Db	5986	TTCCGAATCCAGAAATCAAAGTTGTTTGTCTACTATTTGATCCAAAGCAGTGGGTCTT	6045

QY	5561	GAACTGACAAATAGTGTGCTCGTGTGTTTGGAGTCACTCTTTGTATGAATAAATCTAGTCTTT	5620	QY	6641	ACGACTTATCCCACTGCGCAGCAGCCACTGCTGTAAACAGGATTAGCAGAGCAGGATGTAG	6700
Db	6046	GAACTGACAAATAGTGTGCTCGTGTGTTTGGAGTCACTCTTTGTATGAATAAATCTAGTCTTT	6105	Db	7126	ACGACTTATCCCACTGCGCAGCAGCCACTGCTGTAAACAGGATTAGCAGAGCAGGATGTAG	7185
QY	5621	TGATCTAAATAATCTTGACGAGCAGCAAGCGGATAAATACCCAAATCTAAACTCTTTTAA	5680	QY	6701	CGGTGCTACAGAGTCTTTGAAGTGGTGGCTACTACGGCTACACTAGAAGACAGTAT	6760
Db	6106	TGATCTAAATAATCTTGACGAGCAGCAAGCGGATAAATACCCAAATCTAAACTCTTTTAA	6165	Db	7186	CGGTGCTACAGAGTCTTTGAAGTGGTGGCTACTACGGCTACACTAGAAGACAGTAT	7245
QY	5581	ACGTAAAGGACAAAGTATGCTGCTGTATTAACCCCAATCAGCTGTAGTCTGATC	5740	QY	6761	TTGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT	6820
Db	6166	ACGTAAAGGACAAAGTATGCTGCTGTATTAACCCCAATCAGCTGTAGTCTGATC	6225	Db	7246	TTGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT	7305
QY	5741	CTCATCAACTTGAAGGCACTATCTTCTTTTAGAGAAATTTGGGAGATCGGATATCGAG	5800	QY	6821	CGGCAAAACAAACACCGCTGCTAGCGGTGGTGTGTTTGTTCGAACAGCAGATTAACGC	6880
Db	6226	CTCATCAACTTGAAGGCACTATCTTCTTTTAGAGAAATTTGGGAGATCGGATATCGAG	6285	Db	7306	CGGCAAAACAAACACCGCTGCTAGCGGTGGTGTGTTTGTTCGAACAGCAGATTAACGC	7365
QY	5801	AAAAGGTACGCTGAATTTTAAACGTGAAATTTATCTCAAGATCTCTGCTCGCGCGTTTC	5860	QY	6881	GCAGAAAAAAGGATCTCAAGAGATCCTTTTGATCTTTCTACGGGGTCTGACGCTCAGT	6940
Db	6286	AAAAGGTACGCTGAATTTTAAACGTGAAATTTATCTCAAGATCTCTGCTCGCGCGTTTC	6345	Db	7366	GCAGAAAAAAGGATCTCAAGAGATCCTTTTGATCTTTCTACGGGGTCTGACGCTCAGT	7425
QY	5861	GGTGATCAGCGTGAACCTCTGACATCGAGCTCCCGGAGACGGTCAAGCTTTGCTG	5920	QY	6941	GGAACGAAAACTCAGCTTAAGGATTTTGGTCAATGAGATTATCAAAAAGGATCTTCACT	7000
Db	6346	GGTGATCAGCGTGAACCTCTGACATCGAGCTCCCGGAGACGGTCAAGCTTTGCTG	6405	Db	7426	GGAACGAAAACTCAGCTTAAGGATTTTGGTCAATGAGATTATCAAAAAGGATCTTCACT	7485
QY	5921	TAAGCGATGCGGAGCAGAACAGCCGCTCAGGCGCGCTCAGCGGGTGTGCGGGTGT	5980	QY	7001	AGATCCTTTT	7010
Db	6406	TAAGCGATGCGGAGCAGAACAGCCGCTCAGGCGCGCTCAGCGGGTGTGCGGGTGT	6465	Db	7486	AGATCCTTTTACGCGCCCTGTAGCGGCAATTAAGCGCGGGGTGTGGTGTACGCGC	7545
QY	5981	CGGGCGCAGCATGACCACTACGTACGTAGCGAGTGTATCTGCTTTAACTATG	6040	QY	7011	-----	7010
Db	6466	CGGGCGCAGCATGACCACTACGTACGTAGCGAGTGTATCTGCTTTAACTATG	6525	Db	7546	AGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCGCTCTTTTCGCTTTTCCCTTCC	7605
QY	6041	CGGCATCAGACAGATTGTACTGAGAGTGCAACATATCGGTTGAAATACCGCACAGAT	6100	QY	7011	-----	7010
Db	6526	CGGCATCAGACAGATTGTACTGAGAGTGCAACATATCGGTTGAAATACCGCACAGAT	6585	Db	7606	TTTCTCGCCACGTTGCGCGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGG	7665
QY	6101	CGTAAGAGAGAAATACCGCATCAGGCTCTTCCGCTTCTCGCTCACTGACTCGCTGC	6160	QY	7011	-----	7010
Db	6586	CGTAAGAGAGAAATACCGCATCAGGCTCTTCCGCTTCTCGCTCACTGACTCGCTGC	6645	Db	7666	TTTCTCGCCACGTTGCGCGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGG	7725
QY	6161	GCTCGCTGCTTCCGCTGCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTAT	6220	QY	7011	-----	7010
Db	6646	GCTCGCTGCTTCCGCTGCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTAT	6705	Db	7726	CGTAGTGGGCAATCGCCCTGTATGACGGTCTTCCCGCTTTGACGTTGGAGTCCACGTTT	7785
QY	6221	CCACGAATCAGGGATTAACCGAGAGAAACATGTGAGCAAAAGGCGCAGCAAGGCCA	6280	QY	7011	-----	7010
Db	6706	CCACGAATCAGGGATTAACCGAGAGAAACATGTGAGCAAAAGGCGCAGCAAGGCCA	6765	Db	7786	TTTAAATAGTGGACTCTTGTTCAAACCTGGAACAAACACTCAACCTATCTCGGTCTATTCT	7845
QY	6281	GGAACGTTAAAGCCGCTGCTGCTGCTTTTCCATAGGCTCCGCCCTCCCTTGAAGC	6340	QY	7011	-----	7010
Db	6766	GGAACGTTAAAGCCGCTGCTGCTGCTTTTCCATAGGCTCCGCCCTCCCTTGAAGC	6825	Db	7846	TTTGAATTTAAGGGAATTTCCGATTTCCGCTTTCGGCTTAAATAAATGAGCTGATTTAA	7905
QY	6341	ATCAAAAAATCGAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	6400	QY	7011	-----	7010
Db	6826	ATCAAAAAATCGAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	6885	Db	7906	CAAAAATTTAACCGCAATTTTAAACAAATATTAACGTTTACATTTAAATCAATCTAAAG	7965
QY	6401	AGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCG	6460	QY	7044	TATATATGAGTAAATCTTGGTGTGACAGTCAATGCTTAAATCAAGTGAAGGACCTATCTC	7103
Db	6886	AGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCG	6945	Db	7966	TATATATGAGTAAATCTTGGTGTGACAGTCAATGCTTAAATCAAGTGAAGGACCTATCTC	8025
QY	6461	GATACCTGCGGCTTTCTCCCTTTGGGAAGCGTGGGCTTTCTATAGCTCAGCTGTA	6520	QY	7104	ACGGATCTGCTTATTTGTTTCAATAGTTTGCCTGACTCCCGCTGCTGTAGATAACTAC	7163
Db	6946	GATACCTGCGGCTTTCTCCCTTTGGGAAGCGTGGGCTTTCTATAGCTCAGCTGTA	7005	Db	8026	ACGGATCTGCTTATTTGTTTCAATAGTTTGCCTGACTCCCGCTGCTGTAGATAACTAC	8085
QY	6521	GGTATCTCAGTTCGGTGTAGTGTGCTTCCCTCAAGCTGGGCTGTGCAAGACCCCGG	6580	QY	7164	GATACGGGAGGCTTACCATCTGGCCCCCAGTGTGCAATGATACCGGAGACCCACGCTC	7223
Db	7006	GGTATCTCAGTTCGGTGTAGTGTGCTTCCCTCAAGCTGGGCTGTGCAAGACCCCGG	7065	Db	8086	GATACGGGAGGCTTACCATCTGGCCCCCAGTGTGCAATGATACCGGAGACCCACGCTC	8145
QY	6581	TTACGCGGACCGCTGCGCTTATCCGTTAACTATCGTCTTAACTATCGTCTTAACTAT	6640	QY	7224	ACCGCTCCAGATTTTATCAGCAATTAACAGCCAGCCGAGGGGCGGAGCGCAAGTGG	7283
Db	7066	TTACGCGGACCGCTGCGCTTATCCGTTAACTATCGTCTTAACTATCGTCTTAACTAT	7125	Db	8146	ACCGCTCCAGATTTTATCAGCAATTAACAGCCAGCCGAGGGGCGGAGCGCAAGTGG	8205
				QY	7284	TCCTGCACTTTTATCCGCTTCACTCCAGTCTATTAAATTTGTTCCGCGGAGAGCTAGATGAAG	7343

Db 8206 TCCTGCAACTTATCCGCTCCATCCAGTCTATTAAATGTTGCGGGAAGTAGAGTAAG 8265  
QY 7344 TAGTTCGCCAGTTAATAGTTTGCCCAAGCTGTTGTCATTCGTCAGGCATCGTGGTGTC 7403  
Db 8266 TAGTTCGCCAGTTAATAGTTTGCCCAAGCTGTTGTCATTCGTCAGGCATCGTGGTGTC 8325  
QY 7404 ACGTCGTCGTTGCGTATGCGTTCATTACGCTCCGGTTCCTCCCAAGCATCAAGGGGAGTTAC 7463  
Db 8326 ACGTCGTCGTTGCGTATGCGTTCATTACGCTCCGGTTCCTCCCAAGCATCAAGGGGAGTTAC 8385  
QY 7464 ATGATCCCGCATGTTGTCARAAAGCGGTAGCTCTCGTCTCGATCGTTCGTCAG 7523  
Db 8386 ATGATCCCGCATGTTGTCARAAAGCGGTAGCTCTCGTCTCGATCGTTCGTCAG 8445  
QY 7524 AAGTAAGTTGCGCGCAGTGTATCACTCATGTTATGCGCAGCTGCATATTCCTTAC 7583  
Db 8446 AAGTAAGTTGCGCGCAGTGTATCACTCATGTTATGCGCAGCTGCATATTCCTTAC 8505  
QY 7584 TGTATGCGCATCCGTAAGATGCTTTCTGTCACCTGGTGTAGTACTCAACCAAGTCACTG 7643  
Db 8506 TGTATGCGCATCCGTAAGATGCTTTCTGTCACCTGGTGTAGTACTCAACCAAGTCACTG 8565  
QY 7644 AAGTAAGTTGCGCGCAGTGTATCACTCATGTTATGCGCAGCTGCATATTCCTTAC 7703  
Db 8566 AAGTAAGTTGCGCGCAGTGTATCACTCATGTTATGCGCAGCTGCATATTCCTTAC 8625  
QY 7704 GCCATAGCAGAACTTTAAAGTGTCTCATCTGGAAGAGTCTTCGCGGGGGAAGT 7763  
Db 8626 GCCATAGCAGAACTTTAAAGTGTCTCATCTGGAAGAGTCTTCGCGGGGGAAGT 8685  
QY 7764 CTCAGGATCTTACGCTGTTGATGATCCAGTTCAGTGAACCCACTCGTCACCCCACTG 7823  
Db 8686 CTCAGGATCTTACGCTGTTGATGATCCAGTTCAGTGAACCCACTCGTCACCCCACTG 8745  
QY 7824 ATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAA 7883  
Db 8746 ATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAA 8805  
QY 7884 TGCAGCAAAAGGGAATAAGCGGACACGGAATGTTGATCTCATCTCTCTCTCTTT 7943  
Db 8806 TGCAGCAAAAGGGAATAAGCGGACACGGAATGTTGATCTCATCTCTCTCTCTTT 8865  
QY 7944 TCAATATTATTGAACATTTATCAGGTTATTGCTCATGAGCGGATACATATTGTAATG 8003  
Db 8866 TCAATATTATTGAACATTTATCAGGTTATTGCTCATGAGCGGATACATATTGTAATG 8925  
QY 8004 TATTAGAAAATAAACAATAGGGTTCCGGGACACATTTCCCGAAAAGTGCACCTGA 8063  
Db 8926 TATTAGAAAATAAACAATAGGGTTCCGGGACACATTTCCCGAAAAGTGCACCTGA 8985  
QY 8064 CGTCTAAGAAACCATTTATTATCATGACATTAACCTATAAAATAGGCGTATCAGAGGCC 8123  
Db 8986 CGTCTAAGAAACCATTTATTATCATGACATTAACCTATAAAATAGGCGTATCAGAGGCC 9045  
QY 8124 CTTTCGCTCTCAA-----GAATTAATTCATGTTG 8155  
Db 9046 CTTTCGCTCTCAAAGATTTCTCATGTTTGACAGCTTATCATCGAATTAATTCATGTTG 9105  
QY 8156 ACAGCTTATCATGATAGCTGACTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 8215  
Db 9106 ACAGCTTATCATGATAGCTGACTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 9165  
QY 8216 CACTGAAAAATAACAGTTATTATTCG 8241  
Db 9166 CACTGAAAAATAACAGTTATTATTCG 9191

RESULT 7  
AAD05811  
ID AAD05811 standard; DNA; 9276 BP.  
XX  
AC AAD05811;

31-JUL-2001 (first entry)  
Pichia pastoris vector pPIC9K.  
Endotheliase; protease domain; cytostatic; vulnery; wound; nootropic;  
periodontitis; dermatological disorder; gene therapy; scar; angiogenesis;  
cardiovascular disorder; psoriasis; neovascular disease;  
chronic inflammatory disease; ocular disorder; circulatory disorder;  
liver cirrhosis; atherosclerosis; haemangiomatosis; diabetes mellitus;  
inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;  
systemic vasculitis; scleroderma; neoplasm; ulcer; burn;  
transmembrane serine protease; pPIC9K vector; ds.  
Pichia pastoris.  
XX  
XX Key Location/Qualifiers  
XX promoter 1..948  
XX primer\_bind /tag= a  
/label= AOX1\_promoter\_fragment  
855..875  
XX sig\_peptide /tag= b  
/label= AOX1\_primer\_site  
949..1218  
XX primer\_bind /tag= c  
/notes= Alpha-factor secretion signal  
1152..1172  
XX misc\_feature /tag= d  
/label= Alpha-factor\_primer\_site  
1192..1241  
XX terminator /tag= e  
/label= Multiple\_cloning\_site  
1253..1586  
XX primer\_bind /tag= f  
/notes= "transcription termination region"  
1327..1347  
XX CDS /tag= g  
/label= AOX1 primer site  
complement(4514..1580)  
XX CDS /tag= h  
/product= "HIS4"  
complement(5743..4928)  
XX misc\_feature /tag= i  
/phenotype= "Kanamycin resistance"  
6122..6879  
XX misc\_feature /tag= j  
/label= AOX1 fragment  
complement(7561..7288)  
XX CDS /tag= k  
/label= Cole1 origin  
complement(8966..8106)  
XX (CORV-) CORVAS INT INC.  
/phenotype= "Ampicillin resistance"  
XX WO200136604-A2.  
XX 25-MAY-2001.  
XX 17-NOV-2000; 2000WO-US031803.  
XX 18-NOV-1999; 99US-0166391P.  
XX 22-SEP-2000; 2000US-0234840P.  
XX (CORV-) CORVAS INT INC.  
XX Madison EL, Ong EO;  
XX WPI; 2001-336001/35.  
XX New nucleic acid encoding a protein comprising endotheliase activity  
XX useful in the prevention and treatment of e.g. vascular malformations,  
XX cardiovascular disorders, and chronic inflammatory disease.  
XX

XX

Example 1; Page 146-149; 152pp; English.

CC The present DNA sequence is Pichia pastoris pPIC9K vector. This pPIC9K  
CC plasmid contains AOX1 promoter fragment, AOX1 primer site, alpha-factor  
CC secretion signal, alpha-factor primer site, multiple cloning site, AOX1  
CC transcription termination region, HIS4 open reading frame (ORF), AOX1  
CC fragment, ColE origin, kanamycin and ampicillin resistance gene. The  
CC invention relates to an endothelial protein, endothelial protease  
CC domain and their corresponding nucleic acid molecules. An endothelial  
CC protein or protease domain of it is useful for the treatment and  
CC diagnosis of disorders associated with aberrant angiogenesis or undesired  
CC neovascularisation. The undesired angiogenesis is associated with  
CC disorders selected from solid neoplasm, vascular malformations and  
CC cardiovascular disorders such as angiodioma, angiolipoma,  
CC arteriovenous malformations, haemangiomas and vascular adhesions, dyschondroplasia  
CC with vascular hamartomas (fatucchi's syndrome), hereditary haemorrhagic  
CC telangiectasia (Rendu-Osler-Weber syndrome) and Von Hippel Lindau  
CC syndrome, chronic inflammatory diseases such as diabetes mellitus,  
CC haemophilic joints, inflammatory bowel disease, nonhealing fractures,  
CC pericarditis, psoriasis, rheumatoid arthritis, venous stasis ulcers,  
CC granuloma, burns, hypertrophic scars, liver cirrhosis,  
CC osteoradionecrosis, postoperative adhesion, pyogenic granuloma and  
CC systemic sclerosis and aberrant wound repairs, circulatory disorders  
CC Raynaud's phenomenon, crest syndromes such as calcinosis, oesophageal,  
CC dyomyolysis, sclerodactyly and telegiectasis, dermatological disorders  
CC such as systemic vasculitis, scleroderma, pyoderma gangrenosum,  
CC vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine  
CC stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome  
CC and Osler-Weber-Rendu syndrome and ocular disorders such as blindness  
CC caused by ocular neovascular disease, corneal graft neovascularisation,  
CC macular degeneration, retinopathy of prematurity, retrolental fibroplasia  
CC and corneal neovascularisation. The nucleic acids of the invention are  
CC also used in gene therapy. The invention also provides method for  
CC screening compounds that modulate angiogenesis

SQ Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;

Query Match 79.6%; Score 6561.6; DB 4; Length 9276;

Best Local Similarity 84.7%; Pred. No. 0;

Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY	1	AGATCTACATCCAAAGCAGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAG	60
DB	1	AGATCTACATCCAAAGCAGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAG	60
QY	61	GTCCATCTCTACACATAAGTGCCTCAACGCAACGAGGAGGATACACTAGCAGACCGGT	120
DB	61	GTCCATCTCTACACATAAGTGCCTCAACGCAACGAGGAGGATACACTAGCAGACCGGT	120
QY	121	TGCAAAAGCAGGAGCTCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	180
DB	121	TGCAAAAGCAGGAGCTCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	180
QY	181	AGCCAGGTTATGGGCTTGATGGAGTCTGCTCATTCCTCTCTCTCTCTCTCTCTCTCTCT	240
DB	181	AGCCAGGTTATGGGCTTGATGGAGTCTGCTCATTCCTCTCTCTCTCTCTCTCTCTCTCT	240
QY	241	ACACATGACCTTTATAGCTGTCTATCTGGCCCCCTCGCGAGGTTGATGTTGTTT	300
DB	241	ACACATGACCTTTATAGCTGTCTATCTGGCCCCCTCGCGAGGTTGATGTTGTTT	300
QY	301	TTTCCGATGCAACAGCTCCGATTTACCCGCAATCATCTCCAGATGAGGCTTCTG	360
DB	301	TTTCCGATGCAACAGCTCCGATTTACCCGCAATCATCTCCAGATGAGGCTTCTG	360
QY	361	AGTGTGGGTCAATAGTTTCATGTTCCCAAATGGCCCAAACTGACATTTAAACGCT	420
DB	361	AGTGTGGGTCAATAGTTTCATGTTCCCAAATGGCCCAAACTGACATTTAAACGCT	420
QY	421	GTCTTGGAACTTATATGACAAAGCGTGATCTCATCCAGATGCACTAAGTTTGGTTTCG	480

DB	421	GTCTTGGAACTTATATGACAAAGCGTGATCTCATCCAAAGTGAATAAGTTTGGTTTCG	480
QY	481	TTGAAATCTCTAACCGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCCCATCCGTTTGT	540
DB	481	TTGAAATCTCTAACCGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCCCATCCGTTTGT	540
QY	541	CTTGTTGGTATTAATGATGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGAGTCT	600
DB	541	CTTGTTGGTATTAATGATGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGAGTCT	600
QY	601	CTCTATCGCTTCTCAACCCCGGTGACCTGTGCGGAAACGCAATGGGAAACACCCGCT	660
DB	601	CTCTATCGCTTCTCAACCCCGGTGACCTGTGCGGAAACGCAATGGGAAACACCCGCT	660
QY	661	TTTTGGATGATTAATGATGCTCTCCCAATGTAATGCTTCCAAAGATTTGTTGGGAATCT	720
DB	661	TTTTGGATGATTAATGATGCTCTCCCAATGTAATGCTTCCAAAGATTTGTTGGGAATCT	720
QY	721	GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAAACCTTTTATCATCTT	780
DB	721	GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAAACCTTTTATCATCTT	780
QY	781	ATATAACAGAGGAAGTGTGCTTAAACCTTTTATCATCTTATCATCTTATCATCTT	840
DB	781	ATATAACAGAGGAAGTGTGCTTAAACCTTTTATCATCTTATCATCTTATCATCTT	840
QY	841	ACTTTCATTAATGCGACTGGTTCATGATCAAAATTTAACTGTTCTAAACCTTTTATCATCTT	900
DB	841	ACTTTCATTAATGCGACTGGTTCATGATCAAAATTTAACTGTTCTAAACCTTTTATCATCTT	900
QY	901	CAACTTGAAGAGATCAAAAACAACTAATTTTCCAAAGGATCCAAACGATGAGATTTCT	960
DB	901	CAACTTGAAGAGATCAAAAACAACTAATTTTCCAAAGGATCCAAACGATGAGATTTCT	960
QY	961	TCAATTTTATGCGAGTATTCGCGATCTCCGATCTCCGATCTCCGATCTCCGATCT	1020
DB	961	TCAATTTTATGCGAGTATTCGCGATCTCCGATCTCCGATCTCCGATCTCCGATCT	1020
QY	1021	ACACAGAGATGAACGCGACAAATTCGCGCTGAAGTGTCTCATCGTTTACTCAGATTTA	1080
DB	1021	ACACAGAGATGAACGCGACAAATTCGCGCTGAAGTGTCTCATCGTTTACTCAGATTTA	1080
QY	1081	GAAAGGGATTTTCGATGTTGCTGTTTTCCTAATTTTCCAAAGCACAATAAATACGGTTATTG	1140
DB	1081	GAAAGGGATTTTCGATGTTGCTGTTTTCCTAATTTTCCAAAGCACAATAAATACGGTTATTG	1140
QY	1141	TTTATAATATCTACTATTCGCGATGCTCTCTAAAGAGAGGGGTATCTCTCGAGAA	1200
DB	1141	TTTATAATATCTACTATTCGCGATGCTCTCTAAAGAGAGGGGTATCTCTCGAGAA	1200
QY	1201	AGAGAGGCTGAAGCCCGAGGCCCGAGGCCCGAGGCCCGAGGCCCGAGGCCCGAGGCC	1260
DB	1201	AGAGAGGCTGAAGCCCGAGGCCCGAGGCCCGAGGCCCGAGGCCCGAGGCCCGAGGCC	1260
QY	1261	GCCTCCCGGAGAGGATTAATTTGGCTTCCCGGTGTCCCGGCGAGGAGTGCACGGAG	1320
DB	1261	GCCTCCCGGAGAGGATTAATTTGGCTTCCCGGTGTCCCGGCGAGGAGTGCACGGAG	1320
QY	1321	AGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTTCCCGTGTCTCCACCCCATGGCC	1380
DB	1321	AGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTTCCCGTGTCTCCACCCCATGGCC	1380
QY	1381	ATCGAGAAACATCAAGAGAAAGATGTCCTTTCTAACTAGTGGCGTGAATTCCTTAGG	1440
DB	1381	ATCGAGAAACATCAAGAGAAAGATGTCCTTTCTAACTAGTGGCGTGAATTCCTTAGG	1440
QY	1441	CGGCGCGGATTAATTTCCCTTTAGACATGACTGTTCTCAGTTTGGGCACTTACG	1500
DB	1441	CGGCGCGGATTAATTTCCCTTTAGACATGACTGTTCTCAGTTTGGGCACTTACG	1500
QY	1501	AGAAGACCGGTTCTGCTAGATTTCTTAACAAGAGGATGTCAGAAATGCAATTTGCTGAG	1560
DB	1501	AGAAGACCGGTTCTGCTAGATTTCTTAACAAGAGGATGTCAGAAATGCAATTTGCTGAG	1560
QY	1561	AGAAGACCGGTTCTGCTAGATTTCTTAACAAGAGGATGTCAGAAATGCAATTTGCTGAG	1620
DB	1561	AGAAGACCGGTTCTGCTAGATTTCTTAACAAGAGGATGTCAGAAATGCAATTTGCTGAG	1620

Qy	1561	ATGAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGATTTTTTTT	1620
Db	1355	ATGAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGATTTTTTTT	1414
Qy	1621	GTCAATTTGTTTCTTCGTACGAGCTTGTCTGTGATCAGCTATCTCGAGCTGTAGAA	1680
Db	1415	GTCAATTTGTTTCTTCGTACGAGCTTGTCTGTGATCAGCTATCTCGAGCTGTAGAA	1474
Qy	1681	TATCTGTGTGTAGGGTTTGGGAAATCANTCGAGTTTGATTTTCTTTGGTATTTCCC	1740
Db	1475	TATCTGTGTGTAGGGTTTGGGAAATCANTCGAGTTTGATTTTCTTTGGTATTTCCC	1534
Qy	1741	ACTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTTGGTTGGCAAGCTTATCGATAA	1800
Db	1535	ACTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTTGGTTGGCAAGCTTATCGATAA	1594
Qy	1801	GCTTTAATGCGGTAGTTTATCACAGTTTAAATTTGCTAACGAGTCAAGGACCGTGTATGAA	1860
Db	1595	GCTTTAATGCGGTAGTTTATCACAGTTTAAATTTGCTAACGAGTCAAGGACCGTGTATGAA	1654
Qy	1861	ATCTAACATGCGGTATCTGCTCGGACCGTCAACCTGGATCGTGTAGGATAGG	1920
Db	1655	ATCTAACATGCGGTATCTGCTCGGACCGTCAACCTGGATCGTGTAGGATAGG	1714
Qy	1921	CTTGCTTATGCGGTACTGCGGCGCTCTTCGCGGATATCGTCCATTCGACACATCGC	1980
Db	1715	CTTGCTTATGCGGTACTGCGGCGCTCTTCGCGGATATCGTCCATTCGACACATCGC	1774
Qy	1981	CAGTCACTATGCGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGACCGGT	2040
Db	1775	CAGTCACTATGCGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGACCGGT	1834
Qy	2041	TCTCGGACATGTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTTCGCTACTTGG	2100
Db	1835	TCTCGGACATGTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTTCGCTACTTGG	1894
Qy	2101	AGCCACTATGCACTACGCGATCATGGCGACACACCCCGTCTCTGTGGATCTATCGAATCTA	2160
Db	1895	AGCCACTATGCACTACGCGATCATGGCGACACACCCCGTCTCTGTGGATCTATCGAATCTA	1954
Qy	2161	AATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2220
Db	1955	AATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2014
Qy	2221	CAGCATTCGGGTAGCATCTAGACCTTCAACAGAGCCAGATCCATCACTGCTTGGCCAA	2280
Db	2015	CAGCATTCGGGTAGCATCTAGACCTTCAACAGAGCCAGATCCATCACTGCTTGGCCAA	2074
Qy	2281	TATGTTTCACTGCTCAGAGTTACGTTCTGTGAAGTATGAATCTCTGGAAGTTGCGAG	2340
Db	2075	TATGTTTCACTGCTCAGAGTTACGTTCTGTGAAGTATGAATCTCTGGAAGTTGCGAG	2134
Qy	2341	TGTTAACTCCGCTGATTTGACGGGCATATCCGTAGTTGGCAAGTGTGGTTGTTACCGG	2400
Db	2135	TGTTAACTCCGCTGATTTGACGGGCATATCCGTAGTTGGCAAGTGTGGTTGTTACCGG	2194
Qy	2401	AGGAGTAACTCTCCAACTCTCTGGAGAGTAGGCCAACCAACACACAGATCCAGCTGTT	2460
Db	2195	AGGAGTAACTCTCCAACTCTCTGGAGAGTAGGCCAACCAACACACAGATCCAGCTGTT	2254
Qy	2461	GTACTTGATCAACATTAAGAGAGATTTCTCGATTTGCAAGGATCAAGTGTTCAGGAGCGT	2520
Db	2255	GTACTTGATCAACATTAAGAGAGATTTCTCGATTTGCAAGGATCAAGTGTTCAGGAGCGT	2314
Qy	2521	ACTGATTGACATTTCCAAAGCCTGCTGCTAGTTTGGCAACGATAGGGTTGTAGAGTGTG	2580
Db	2315	ACTGATTGACATTTCCAAAGCCTGCTGCTAGTTTGGCAACGATAGGGTTGTAGAGTGTG	2374
Qy	2581	CAATACACTTGGGTCAATTTCAAACCCCTTGGCAACTGCAAGCTTGGTTGTAACAGCAT	2640
Db	2375	CAATACACTTGGGTCAATTTCAAACCCCTTGGCAACTGCAAGCTTGGTTGTAACAGCAT	2434

Qy	2641	CTTCAATTTCTGGCAAGCTTCTTGTCTGTATATCGACAGCCAAAGATACCTCGGAAAT	2700
Db	2435	CTTCAATTTCTGGCAAGCTTCTTGTCTGTATATCGACAGCCAAAGATACCTCGGAAAT	2494
Qy	2701	CAATACCATTTTCTGAGTTCAGAGTCTCTGAGGCAACGAAATCTGGATCAGCGTATTT	2760
Db	2495	CAATACCATTTTCTGAGTTCAGAGTCTCTGAGGCAACGAAATCTGGATCAGCGTATTT	2554
Qy	2761	TATCAGCATATTAAGAACTTTCAGAGGCCCGCAGCAGGATGTCAATCTACACAGGCTG	2820
Db	2555	TATCAGCATATTAAGAACTTTCAGAGGCCCGCAGCAGGATGTCAATCTACACAGGCTG	2614
Qy	2821	ATGTGTCAATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTCGACCAAAATA	2880
Db	2615	ATGTGTCAATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTCGACCAAAATA	2674
Qy	2881	TTTTGTCAACATTAAGAACTTCTGTTTCCGTAAAGCCATAGCAGCTACTGCTGGGCGC	2940
Db	2675	TTTTGTCAACATTAAGAACTTCTGTTTCCGTAAAGCCATAGCAGCTACTGCTGGGCGC	2734
Qy	2941	CTTCCTGTAGCAGTATACACTTTAGCACCAACTTGTGGCAACGATAGATGATTTCTGGGG	3000
Db	2735	CTTCCTGTAGCAGTATACACTTTAGCACCAACTTGTGGCAACGATAGATGATTTCTGGGG	2794
Qy	3001	TAAGGTACCATCTCTTTAGTGGAGATGCAAAAACAAATTTCTTTGCAACAGCAACTTT	3060
Db	2795	TAAGGTACCATCTCTTTAGTGGAGATGCAAAAACAAATTTCTTTGCAACAGCAACTTT	2854
Qy	3061	TGSCAGGACACCCAGCATCAGGAGTGTGAAGGAGAGATTCGGGTTCCACAGCAATAT	3120
Db	2855	TGSCAGGACACCCAGCATCAGGAGTGTGAAGGAGAGATTCGGGTTCCACAGCAATAT	2914
Qy	3121	AGAGGCCAACTTTCTCAATAGTCTTGTCAAAACGAGAGCAGACTACACAGGGCGAAGTCT	3180
Db	2915	AGAGGCCAACTTTCTCAATAGTCTTGTCAAAACGAGAGCAGACTACACAGGGCGAAGTCT	2974
Qy	3181	CAACTTGGCAACCGTCTCCGTTAGTGTGAGCTTCTATGGAATTTCTCTGAGAGTATCTATAGAG	3240
Db	2975	CAACTTGGCAACCGTCTCCGTTAGTGTGAGCTTCTATGGAATTTCTCTGAGAGTATCTATAGAG	3034
Qy	3241	GATCAATGCTCTCTTAAAGTATCTGCAATTTGCAATTAAGTTTCTCTGGAAAGGAGCTT	3300
Db	3035	GATCAATGCTCTCTTAAAGTATCTGCAATTTGCAATTAAGTTTCTCTGGAAAGGAGCTT	3094
Qy	3301	CTAACACAGGTGTCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTTT	3360
Db	3095	CTAACACAGGTGTCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTTT	3154
Qy	3361	TGTCACCATTTTGAACGACATTTGACAAATTTGACTTATCCATAATCTGTTCCG	3420
Db	3155	TGTCACCATTTTGAACGACATTTGACAAATTTGACTTATCCATAATCTGTTCCG	3214
Qy	3421	TTTTCTGGATAGGACGACGAGGCGATCTTCAATTTCTTGTGAGGAGGCTTTAGAAACGT	3480
Db	3215	TTTTCTGGATAGGACGACGAGGCGATCTTCAATTTCTTGTGAGGAGGCTTTAGAAACGT	3274
Qy	3481	CAATTTTGGCAATTTCAATACGACTTTCAGAGGAGCTTCTTTAGGTTTGGATTTCTTCT	3540
Db	3275	CAATTTTGGCAATTTCAATACGACTTTCAGAGGAGCTTCTTTAGGTTTGGATTTCTTCT	3334
Qy	3541	TAGTGTGTTCTTGGTGTATCTTGGCTTGGCATCTCTTTCTTCTAGTGAACCTTTAGGG	3600
Db	3335	TAGTGTGTTCTTGGTGTATCTTGGCTTGGCATCTCTTTCTTCTAGTGAACCTTTAGGG	3394
Qy	3601	ACTTCATATCCAGGTTTCTCTCACCTGTCGCAACGATCAACGATCTTGGGCACTTAA	3660
Db	3395	ACTTCATATCCAGGTTTCTCTCACCTGTCGCAACGATCAACGATCTTGGGCACTTAA	3454
Qy	3661	CTAATCCAAATAAATAAGTACAGCACTTCCAGGCTATATCTTCTTCTTGGATTTAGCTT	3720
Db	3455	CTAATCCAAATAAATAAAGTACAGCACTTCCAGGCTATATCTTCTTCTTGGATTTAGCTT	3514
Qy	3721	CTGCAAGTTTCAATCAGCTTCTCTCCCTAAATTTTAGCGTTCAACAAACCTCTGCTGCAATA	3780

Db 3515 CTGCAAGTTCAACAGCTTCTCCCTAATTTAGGTTCAACAAACTTTCGTCTCAANTA 3574  
Qy 3781 ACCGTTGGTATAAGAACCTTCTGGAGATGTGCTTTACGATCCCAAGGTGGCTTCCA 3840  
Db 3575 ACCGTTGGTATAAGAACCTTCTGGAGATGTGCTTTACGATCCCAAGGTGGCTTCCA 3634  
Qy 3841 TGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAAGTGGCTTCAAGTGACAGAAACCAA 3900  
Db 3635 TGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAAGTGGCTTCAAGTGACAGAAACCAA 3694  
Qy 3901 CACCTGTTTGTTCACACCAAAATTCAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3960  
Db 3695 CACCTGTTTGTTCACACCAAAATTCAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3754  
Qy 3961 GCAACTTTTGAGTTGCTCCAGATGTAGCACTTTATACCAACCGTGACGAGGATT 4020  
Db 3755 GCAACTTTTGAGTTGCTCCAGATGTAGCACTTTATACCAACCGTGACGAGGATT 3814  
Qy 4021 GGTAGACTCCAGTTTGTGCTTATAGCTCCCGAATAGACTTTTGGACGAGTACACCA 4080  
Db 3815 GGTAGACTCCAGTTTGTGCTTATAGCTCCCGAATAGACTTTTGGACGAGTACACCA 3874  
Qy 4081 GGCCCAACGAGTAATTAGAGAGTCAGCCACCAAAAGTAGTGAATAGACCAATCGGGGGGT 4140  
Db 3875 GGCCCAACGAGTAATTAGAGAGTCAGCCACCAAAAGTAGTGAATAGACCAATCGGGGGGT 3934  
Qy 4141 CAGTAGTCAAAAGAGCCAAACAAATTTCACTGACAGGAACTTTTGACATCTTCAGAAA 4200  
Db 3935 CAGTAGTCAAAAGAGCCAAACAAATTTCACTGACAGGAACTTTTGACATCTTCAGAAA 3994  
Qy 4201 GTTCGTAATTCAGTAGTCAANTGCGAGATCAATAATGGGGAATTATACCAAGAGCAACAG 4260  
Db 3995 GTTCGTAATTCAGTAGTCAANTGCGAGATCAATAATGGGGAATTATACCAAGAGCAACAG 4054  
Qy 4261 TGGAGTCAACATCTACCAACTTTTCGGTCTCAGAAAAGATATAACAGTTCTACTACCGC 4320  
Db 4055 TGGAGTCAACATCTACCAACTTTTCGGTCTCAGAAAAGATATAACAGTTCTACTACCGC 4114  
Qy 4321 CATTAGTGAACCTTTTCAATTCGCCAATCTAGGTCACAAATCACTTCATTGATACCAT 4380  
Db 4115 CATTAGTGAACCTTTTCAATTCGCCAATCTAGGTCACAAATCACTTCATTGATACCAT 4174  
Qy 4381 TAGGGGGAAGAGTCAACCTTTTATCAACAGGTCCTATAGATTAACCTAGCGCTGGGA 4440  
Db 4175 TAGGGGGAAGAGTCAACCTTTTATCAACAGGTCCTATAGATTAACCTAGCGCTGGGA 4234  
Qy 4441 TCATCTTTTGGACAACTTTTCTGCCAATCTAGGTCACAAATCACTTCATTGATACCAT 4500  
Db 4235 TCATCTTTTGGACAACTTTTCTGCCAATCTAGGTCACAAATCACTTCATTGATACCAT 4294  
Qy 4501 TATTGTACAACTTGAGCAAGTTGCGATCAGCTCCTCAAAATGGTCTCTGTACCGATG 4560  
Db 4295 TATTGTACAACTTGAGCAAGTTGCGATCAGCTCCTCAAAATGGTCTCTGTACCGATG 4354  
Qy 4561 ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTTGAGTCAACTTGCATCAGGTTGTGCA 4620  
Db 4355 ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTTGAGTCAACTTGCATCAGGTTGTGCA 4414  
Qy 4621 GCTGGTCAGCAGTATAGGGAACACGGCTTTTCTCAACAACTCAAGGAATTTATCAAAC 4680  
Db 4415 GCTGGTCAGCAGTATAGGGAACACGGCTTTTCTCAACAACTCAAGGAATTTATCAAAC 4474  
Qy 4681 CTGCAACATTTGCGTATGTCAGTAGCAGGGAATGTCACTTGAAGTCGGACAGTGAG 4740  
Db 4475 CTGCAACATTTGCGTATGTCAGTAGCAGGGAATGTCACTTGAAGTCGGACAGTGAG 4534  
Qy 4741 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATATCAGTGAAGTCAGTCATCAGGAGA 4800  
Db 4535 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATATCAGTGAAGTCAGTCATCAGGAGA 4594  
Qy 4801 TCCTCTACGCGGAGCATCTGCGCGA----- 4828

Db 4595 TCCTCTACGCGGAGCATCTGTCGCGACATCTCAGGGGGGGGGGGGGGCTGAGGTCTGC 4654  
Qy 4829----- 4828  
Db 4655 CTCGTGAAGAAGGTGTTGCTGACTATACAGGCTGAATCGCCCCATCATCCAGCCAGA 4714  
Qy 4829----- 4828  
Db 4715 AAGTCAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGTGGAACCACTGATTTTGA 4774  
Qy 4829----- 4828  
Db 4775 ACTTTTGTTCGCCACGGAACGGTCTGCGTTGTCGGGAAGATGCGTATCTGATCTTCA 4834  
Qy 4829----- 4828  
Db 4835 ACTCAGCAAAAGTTCGATTTATTCAACAAAGCCGCTCCGTCAGTCAAGTCAGCGTAATGCT 4894  
Qy 4829----- 4828  
Db 4895 CTGCCAGTGTACAAACCAATTACCAATTTCTGATTAGAAAAACTCATCGAGCATCAAAATG 4954  
Qy 4829----- 4828  
Db 4955 AAATGCAATTTATTCATATCAGGATTAATCAATACCATATTTTGAACAAAGCCGTTTCTG 5014  
Qy 4829----- 4828  
Db 5015 TAATGAAGGAGAAAACTCACCGAGCAGTTCATAGGATGGCAAGATCCTGGTATCGGTC 5074  
Qy 4829----- 4828  
Db 5075 TCGGATTCGGACTCGTCCAAACATCAATACACCTATTAAATTTCCCTTCGTCAAAAATAAG 5134  
Qy 4829----- 4828  
Db 5135 GTTATCAAGTGAGAAATCACCATGATGACGACTGAATCCGTCGAGAAATGGCAAAAGCTT 5194  
Qy 4829----- 4828  
Db 5195 ATGCATTTCTTTCCAGACTTGTTCACAGGCGCCATTACGCTGTCATCAAAATCACT 5254  
Qy 4829----- 4828  
Db 5255 CGCATCAACCAACCGTTATTTCATTTCGTGATTGCGCTGAGCGAGACGAATAACGCGATC 5314  
Qy 4829----- 4828  
Db 5315 GCTGTTAAAGGACAATTAACAAAGGAATCGAATGCAACCGCGCGAGGAACACTGCCAG 5374  
Qy 4829----- 4828  
Db 5375 CCATCAACAAATATTTTCACTGATCAGGATATTTCTTAATACCTGGAATGCTGTTTT 5434  
Qy 4829----- 4828  
Db 5435 CCGGGGATCGCAGTGGTGAATAACCATGTCATCAGGATACGGATAAAATGCTTGAAT 5494  
Qy 4829----- 4828  
Db 5495 GGTGGAAGAGGCATAAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAAACATC 5554  
Qy 4829----- 4828  
Db 5555 ATTGCAACGCTACCTTTTCCCATGTTTCAGAAACAACTCTGGCGCATCGGCTTCCCATATA 5614  
Qy 4829----- 4828  
Db 5615 CAATCGATAGATTGTCGCACTGATTGCGGACATTTATCGGAGCCCATTTTATACCCATA 5674  
Qy 4829----- 4828  
Db 5675 TAAATCAGCATCCAGTTTGAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5734

QY 4829 ----- 4828  
Db 5735 ATGGCTATAACACCCCTTGTAATCTACTGTTATGTAAACACACAGTATTTATTTGTTCAAGA 5794  
QY 4829 ----- 4828  
Db 5795 TGATATATTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTCGGCTTTC 5854  
QY 4829 ----- 4828  
Db 5855 CCGGCTGCGGCTGCGAGGTCGGATCACCAGGCGCACAGTGGGTTGCTGGCGCTATAT 5914  
QY 4880 CGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATAGCGCTTGTT 4939  
Db 5915 CGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATAGCGCTTGTT 5974  
QY 4940 CGGCGTGGGTATGTTGGCAGGCGGCTGGCGGGGACTGTTGGCGGCATCTCTTTGCA 4999  
Db 5975 CGGCGTGGGTATGTTGGCAGGCGGCTGGCGGGGACTGTTGGCGGCATCTCTTTGCA 6034  
QY 5000 TGCACCATTCCTTGGCGGGGGTCTCAACGGGCTCAACCTACTACTTGGGCTGCTTCT 5059  
Db 6035 TGCACCATTCCTTGGCGGGGGTCTCAACGGGCTCAACCTACTACTTGGGCTGCTTCT 6094  
QY 5060 AATGAGAGTCGCATAAGGAGAGCGTCGAGTATCTATGATTTGAAGTATGGGAATGGT 5119  
Db 6095 AATGAGAGTCGCATAAGGAGAGCGTCGAGTATCTATGATTTGAAGTATGGGAATGGT 6154  
QY 5120 GATACCCGATTCCTCAGTGTCTTCAGGTCTCCATCAGATTTAGCCCACTAAAGCAAC 5179  
Db 6155 GATACCCGATTCCTCAGTGTCTTCAGGTCTCCATCAGATTTAGCCCACTAAAGCAAC 6214  
QY 5180 CGGAGGAGGATTCATGGTAAATTCCTGACTTTTGGTCATCAGTAGACTCGAACTG 5239  
Db 6215 CGGAGGAGGATTCATGGTAAATTCCTGACTTTTGGTCATCAGTAGACTCGAACTG 6274  
QY 5240 TGAGACTATCTCGGTATGACAGCAGAAATGTCCTTCTTGAGACAGTAAATGAGTCCC 5299  
Db 6275 TGAGACTATCTCGGTATGACAGCAGAAATGTCCTTCTTGAGACAGTAAATGAGTCCC 6334  
QY 5300 ACCAATAAGAAATCCTTTGTTATCAGGAACAAACTTCTTTGAACTTTTCGGTGCC 5359  
Db 6335 ACCAATAAGAAATCCTTTGTTATCAGGAACAAACTTCTTTGAACTTTTCGGTGCC 6394  
QY 5360 TTGAACATAAATATAGAGTGGAATGTCGGGTAGGAATGAGCGGCAAAATGCTTACC 5419  
Db 6395 TTGAACATAAATATAGAGTGGAATGTCGGGTAGGAATGAGCGGCAAAATGCTTACC 6454  
QY 5420 TTCTCGACCTTCAAGAGGTATCTAGGGTTTGTAGATACTGATGCCAATTCAGTGACAAC 5479  
Db 6455 TTCTCGACCTTCAAGAGGTATCTAGGGTTTGTAGATACTGATGCCAATTCAGTGACAAC 6514  
QY 5480 GTTGCTATTTCTTCAACCAATTCGGAATCCAGAGAAATCAAGTGTGTTGCTACTATT 5539  
Db 6515 GTTGCTATTTCTTCAACCAATTCGGAATCCAGAGAAATCAAGTGTGTTGCTACTATT 6574  
QY 5540 GATCAAGCCGTCGGCTTGAACTGACATAGTGTCTCGTGTGTTTGAAGTCACTTT 5599  
Db 6575 GATCAAGCCGTCGGCTTGAACTGACATAGTGTGCTCGTGTGTTTGAAGTCACTTT 6634  
QY 5600 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTCACGAGCAAGCGCATAAATACC 5659  
Db 6635 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTCACGAGCAAGCGCATAAATACC 6694  
QY 5660 CAAATCTAAACTCTTTTAAACGTTTAAAGGACAAAGTATGTCGCCGTGTTTAAACCCC 5719  
Db 6695 CAAATCTAAACTCTTTTAAACGTTTAAAGGACAAAGTATGTCGCCGTGTTTAAACCCC 6754  
QY 5720 AAATCAGCTCGTAGTCTGATCCTCATCACTTGAAGGGCACTATCTCTTTTGAAGAAAT 5779  
Db 6755 AAATCAGCTCGTAGTCTGATCCTCATCACTTGAAGGGCACTATCTCTTTTGAAGAAAT 6814

QY 5780 TTGCGAGATGCGATATCGAGAAAAAGGTACGCTGATTTTAAACGTGAATTTATCTCAA 5839  
Db 5815 TTGCGAGATGCGATATCGAGAAAAAGGTACGCTGATTTTAAACGTGAATTTATCTCAA 5874  
QY 5840 GATCTCTGCTCGCGGTTTGGTGATGACGGTGAATAAACCCTCTGAACATGACAGTCCCG 5899  
Db 5875 GATCTCTGCTCGCGGTTTGGTGATGACGGTGAATAAACCCTCTGAACATGACAGTCCCG 5934  
QY 5900 GACACGCTCACAGCTTGTCTGTAAGCGGATCCGGGAGCAGACAAGCCGTCAGGGCGCG 5959  
Db 5935 GACACGCTCACAGCTTGTCTGTAAGCGGATCCGGGAGCAGACAAGCCGTCAGGGCGCG 5994  
QY 5960 TCAGCGGCTGTTGGCGGCTGTCGGGCGCAGCGCATACCCAGTCACTAGCGATACCGGA 6019  
Db 5995 TCAGCGGCTGTTGGCGGCTGTCGGGCGCAGCGCATACCCAGTCACTAGCGATACCGGA 7054  
QY 6020 GTCTATCTCGCTTAATCTACTATGCGGCATCAGACAGATTTGCTGAGTGCACCATATGC 6079  
Db 7055 GTCTATCTCGCTTAATCTACTATGCGGCATCAGACAGATTTGCTGAGTGCACCATATGC 7114  
QY 6080 GGTGTGAATATACCGCATGATGCTAAGGAGAAAAATACCGCATCAGCGCTCTTTCGCTT 6139  
Db 7115 GGTGTGAATATACCGCATGATGCTAAGGAGAAAAATACCGCATCAGCGCTCTTTCGCTT 7174  
QY 6140 CTTGCTCTACTGACTGCTGCGCTCGGTCTGCGCTGCGGCGAGCGGTATCAGCTCACT 6199  
Db 7175 CTTGCTCTACTGACTGCTGCGCTCGGTCTGCGCTGCGGCGAGCGGTATCAGCTCACT 7234  
QY 6200 CAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACCGAGAAAGAACATGTAG 6259  
Db 7235 CAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACCGAGAAAGAACATGTAG 7294  
QY 6260 CAAAAGCGCAGAAAAAGGCGAGAACCGTAAAGGCGCGGTTGCTGGCGTTTTTCCATA 6319  
Db 7295 CAAAAGCGCAGAAAAAGGCGAGAACCGTAAAGGCGCGGTTGCTGGCGTTTTTCCATA 7354  
QY 6320 GGCTCGCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGCGAAACC 6379  
Db 7355 GGCTCGCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGCGAAACC 7414  
QY 6380 CGACAGACTATTAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 6439  
Db 7415 CGACAGACTATTAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 7474  
QY 6440 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAGCGTGGCG 6499  
Db 7475 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAGCGTGGCG 7534  
QY 6500 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTTAGGTCTGCTCCAAAGCTG 6559  
Db 7535 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTTAGGTCTGCTCCAAAGCTG 7594  
QY 6560 GCTGTGTGACAGAACCCCGCTTACGCGGACCGCTGCGGCTTATCCGGTAACTATCGTC 6619  
Db 7595 GCTGTGTGACAGAACCCCGCTTACGCGGACCGCTGCGGCTTATCCGGTAACTATCGTC 7654  
QY 6620 TTCAGTCCAAACCGGTGAAGACAGCTTATCGCACTGCGCAGCAGCACTGGTAACAGGA 6679  
Db 7655 TTGAGTCCAAACCGGTGAAGACAGCTTATCGCACTGCGCAGCAGCACTGGTAACAGGA 7714  
QY 6680 TTAGCAGACGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACG 6739  
Db 7715 TTAGCAGACGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACG 7774  
QY 6740 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTGCTGAGCCAGTCACTTCCTCGGA 6799  
Db 7775 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTGCTGAGCCAGTCACTTCCTCGGA 7834  
QY 6800 AAAGAGTTGTTAGTCTTGTATCCGCAAAACAAACCAACCGCTGCTAGCGGTGGTTTTTG 6859  
Db 7835 AAAGAGTTGTTAGTCTTGTATCCGCAAAACAAACCAACCGCTGCTAGCGGTGGTTTTTG 7894  
QY 6860 TTTGCAAGCAGCAGATTACGCGCAAAAAAAGGATCTCAAGAAAGATCTCTTGATCTTTT 6919



Db	7895	TTTTCAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTT	7954	Db	8975	TTTTCAATATTATTGAAGCATTATCAGGTTATTGTCTCATGAGCGGATACATATTTG	9034
Qy	6920	CTACGGGCTGAGCGCTCAGTGGAAAGCAAACTCAGCTTAAGGATTTTGGTTCATGAGAT	6979	Qy	8000	AATGTTATTGAAAAATAAATAAGGTTTCGGCAGACATTTCCCGAAAAGTGGCAC	8059
Db	7955	CTACGGGCTGAGCGCTCAGTGGAAAGCAAACTCAGCTTAAGGATTTTGGTTCATGAGAT	8014	Db	9035	AATGTTATTGAAAAATAAATAAGGTTTCGGCAGACATTTCCCGAAAAGTGGCAC	9094
Qy	6980	TATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAATCT	7039	Qy	8060	CTGAGCTCTAAGAACCAATTTATCATGACATTAACTATAAAATAGGCGTATCAGA	8119
Db	8015	TATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAATCT	8074	Db	9095	CTGAGCTCTAAGAACCAATTTATCATGACATTAACTATAAAATAGGCGTATCAGA	9154
Qy	7040	AAAGTATATAGTAATAGTCTGACAGTTACCAATGCTTAATCAAGTGGGACCTTA	7099	Qy	8120	GGCCCTTTTCCTTCAAGAAATTAATTCATGTTGACAGCTTATCATCGAATAGCTGAC	8179
Db	8075	AAAGTATATAGTAATAGTCTGACAGTTACCAATGCTTAATCAAGTGGGACCTTA	8134	Db	9155	GGCCCTTTTCCTTCAAGAAATTAATTCATGTTGACAGCTTATCATCGAATAGCTGAC	9214
Qy	7100	TCTCAGGATCTGTCTATTTCCTTATCCATAGTTGCTGACTCCCGCTGCTAGATAA	7159	Qy	8180	TCATGTTGGTATTGTAATAGACGAGATCGGAAACACTGAAAAATAACAGTTATTATT	8239
Db	8135	TCTCAGGATCTGTCTATTTCCTTATCCATAGTTGCTGACTCCCGCTGCTAGATAA	8194	Db	9215	TCATGTTGGTATTGTAATAGACGAGATCGGAAACACTGAAAAATAACAGTTATTATT	9274
Qy	7160	CTACGATACGGGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGGAGCCAC	7219	Qy	8240	CG 8241	
Db	8195	CTACGATACGGGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGGAGCCAC	8254	Db	9275	CG 9276	
Qy	7220	GCTCACCGGCTCCAGATTATACAGCAATAAACACGAGCGGAGGCGGAGCGAGAA	7279				
Db	8255	GCTCACCGGCTCCAGATTATACAGCAATAAACACGAGCGGAGGCGGAGCGAGAA	8314				
Qy	7280	GTGTCCTTCGAACCTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAAGCTAGAG	7339				
Db	8315	GTGTCCTTCGAACCTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAAGCTAGAG	8374				
Qy	7340	TAAGTAGTCCCGAGTTAATAGTTTGGCAACGTTGTCATTCGTCGAGGATCTGG	7399				
Db	8375	TAAGTAGTCCCGAGTTAATAGTTTGGCAACGTTGTCATTCGTCGAGGATCTGG	8434				
Qy	7400	TGTCAAGCTCGTCTGTTGTTGATGCTTCATTCAGCTCCGTTCCCAACGATCAAGGCGAG	7459				
Db	8435	TGTCAAGCTCGTCTGTTGTTGATGCTTCATTCAGCTCCGTTCCCAACGATCAAGGCGAG	8494				
Qy	7460	TTACATGATCCCGATGTTTGGCAAAAGGGTGTAGTCTCTTCGTCCTCCGATCTTG	7519				
Db	8495	TTACATGATCCCGATGTTTGGCAAAAGGGTGTAGTCTCTTCGTCCTCCGATCTTG	8554				
Qy	7520	TCAGAAAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACTGCATAATCTC	7579				
Db	8555	TCAGAAAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACTGCATAATCTC	8614				
Qy	7580	TTACTGTATGCGCATCCGTAAGATGCTTTCTGTGACTGTGTGACTTCAACCAAGTCAT	7639				
Db	8615	TTACTGTATGCGCATCCGTAAGATGCTTTCTGTGACTGTGTGACTTCAACCAAGTCAT	8674				
Qy	7640	TCTGAGATAGTGTATGCGGAGCGAGTGTCTCTTGGCCGCGCTCAACACGGGATAATA	7699				
Db	8675	TCTGAGATAGTGTATGCGGAGCGAGTGTCTCTTGGCCGCGCTCAACACGGGATAATA	8734				
Qy	7700	CGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGAAAAAGCTTTCTTGGGGCGAA	7759				
Db	8735	CGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGAAAAAGCTTTCTTGGGGCGAA	8794				
Qy	7760	AACTCTCAGGATCTTACCGTGTGTGATGATCCAGTTCGATGTAACCACTGTCACCCA	7819				
Db	8795	AACTCTCAGGATCTTACCGTGTGTGATGATCCAGTTCGATGTAACCACTGTCACCCA	8854				
Qy	7820	ACTGATCTTCAAGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAGGC	7879				
Db	8855	ACTGATCTTCAAGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAGGC	8914				
Qy	7880	AAATGCCGCAAAAAGGATTAAGGCGACACGGAATGTGTATCTACTATCTCTTCC	7939				
Db	8915	AAATGCCGCAAAAAGGATTAAGGCGACACGGAATGTGTATCTACTATCTCTTCC	8974				
Qy	7940	TTTTTCAATATTATTGAAGCATTATATCAGGTTATTGTCTCATGAGCGGATACATATTTG	7999				

RESULT 8  
AAD13151  
ID AAD13151 standard; DNA; 9276 BP.  
XX  
AC AAD13151;  
XX  
DT 16-OCT-2001 (first entry)  
XX  
DE Pichia pastoris pPIC9K vector DNA.  
XX  
KW Human; transmembrane serine protease; membrane-type serine protease;  
KW MTSP; protease domain; neoplastic disease; tumour; cancer; cytostatic;  
KW lung carcinoma; colon adenocarcinoma; ovarian carcinoma; gene therapy;  
KW cyclic; circular; ds.  
XX  
OS Pichia pastoris.  
XX  
PN WO200157194-A2.  
XX  
XX 09-AUG-2001.  
PD  
XX 02-FEB-2001; 2001WO-US003471.  
XX  
XX 03-FEB-2000; 2000US-0179982P.  
XX  
XX 18-FEB-2000; 2000US-0183542P.  
PR  
XX 22-JUN-2000; 2000US-0213124P.  
PR  
XX 26-JUL-2000; 2000US-0220970P.  
PR  
XX 08-SEP-2000; 2000US-00657986.  
PR  
XX 22-SEP-2000; 2000US-0234840P.  
XX  
XX (CORV-) CORVAS INT INC.  
XX  
XX Madison EL, Ong EO, Yeh J;  
PI  
XX MPI; 2001-488877/53.  
DR  
XX  
XX Novel single chain polypeptide comprising protease domain of type-II  
PT membrane-type serine protease or its catalytically active portion useful  
PT for treating and preventing cancer and tumor.  
XX  
XX Example 4; Page 221-223; 256pp; English.  
PS  
XX The invention relates to transmembrane serine proteases and their  
CC corresponding nucleotides and the protease domain of a type-II membrane-  
CC type serine protease (MTSP). MTSP is useful for identifying compounds  
CC that modulate or inhibit its proteolytic activity and for formulating a  
CC medicament for treating neoplastic disease. MTSP and its corresponding  
CC nucleotides are useful in preventing or treating tumours or cancers such  
CC as lung carcinoma, colon adenocarcinoma and ovarian carcinoma, in  
CC diagnostics and in hybridisation assays. MTSP is useful as a diagnostic

CC marker for tumour development, growth and/or progression and as  
CC immunogens to generate antibodies that specifically bind to it. MTSP DNA  
CC is useful in a yeast two-hybrid system and in gene therapy. The present  
CC sequence is Pichia pastoris pPIC9K vector DNA  
xx  
SQ Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;

Query Match 79.6%; Score 6561.6; DB 4; Length 9276;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY 1 AGATCTAATCCAAAGAGCAAGAGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60  
DB 1 AGATCTAATCCAAAGAGCAAGAGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60  
QY 61 GTCCATTCTCACATAAGTSCAAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120  
DB 61 GTCCATTCTCACATAAGTSCAAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120  
QY 121 TGCACAGCAGGACTCCACT 180  
DB 121 TGCACAGCAGGACTCCACT 180  
QY 181 AGCCGAGTTATGGGCTTGATTTGAGCTCGCTCATTCCTTCTATTAGGCTACTA 240  
DB 181 AGCCGAGTTATGGGCTTGATTTGAGCTCGCTCATTCCTTCTATTAGGCTACTA 240  
QY 241 ACACATGACTTTATAGCTGTCTATCTGGGCCCCCTGGCGAGGTTCTATGTTGTTA 300  
DB 241 ACACATGACTTTATAGCTGTCTATCTGGGCCCCCTGGCGAGGTTCTATGTTGTTA 300  
QY 301 TTTCGGAATCAACAAAGCTCGCATACACCCGCAACATCACTCCAGATGAGGGCTTCTG 360  
DB 301 TTTCGGAATCAACAAAGCTCGCATACACCCGCAACATCACTCCAGATGAGGGCTTCTG 360  
QY 361 AGTGCGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCAAAACAGTGAAGTTAAACGCT 420  
DB 361 AGTGCGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCAAAACAGTGAAGTTAAACGCT 420  
QY 421 GTCCTGGAACCTAATATGACAAAGCGTGATCTCATCCAGATGAACTAAGTTTGGTTCG 480  
DB 421 GTCCTGGAACCTAATATGACAAAGCGTGATCTCATCCAGATGAACTAAGTTTGGTTCG 480  
QY 481 TTGAAATGCTAACGGCCAGTTGGTCAAAAGAGAACTTCCAAAAGTCGCCATACCGTTGT 540  
DB 481 TTGAAATGCTAACGGCCAGTTGGTCAAAAGAGAACTTCCAAAAGTCGCCATACCGTTGT 540  
QY 541 CTGTTGGTATTTGATGACGAAATGCTCAAAATAATCTCATTAATGCTTAGCGAGTCT 600  
DB 541 CTGTTGGTATTTGATGACGAAATGCTCAAAATAATCTCATTAATGCTTAGCGAGTCT 600  
QY 601 CTCATCGCTTCTGAACCCCGGTGACCTGTGCCGAAACGCAATGGGAAACACCCGCT 660  
DB 601 CTCATCGCTTCTGAACCCCGGTGACCTGTGCCGAAACGCAATGGGAAACACCCGCT 660  
QY 661 TTTTGGATGATTATGCAATGTCTCCACATTTGATGTTTCCAAAGATTTCTGTTGGAACT 720  
DB 661 TTTTGGATGATTATGCAATGTCTCCACATTTGATGTTTCCAAAGATTTCTGTTGGAACT 720  
QY 721 GCTGATAGCTAAACGCTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAAT 780  
DB 721 GCTGATAGCTAAACGCTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAAT 780  
QY 781 ATATAAAGAGAGGAGTGGCTCTCTTAAACCTTTTTTTTATCATCATTTATTAGCTT 840  
DB 781 ATATAAAGAGAGGAGTGGCTCTCTTAAACCTTTTTTTTATCATCATTTATTAGCTT 840  
QY 841 ACTTTCATAATTTGGACCTGGTTCGAATGACAGCTTTTGATTTTAAGCACTTTTAACGA 900  
DB 841 ACTTTCATAATTTGGACCTGGTTCGAATTTGAATTTTAAGCACTTTTAACGA 900  
QY 901 CAACCTTGAGAGATCAAAAACAACTAATTTTCGAGGATCCAAAACGATCAGATTTTCT 960  
DB 901 CAACCTTGAGAGATCAAAAACAACTAATTTTCGAGGATCCAAAACGATCAGATTTTCT 960

DB 901 CAACTTGAGAGATCAAAAACAACTAATTTTCGAGGATCCAAAACGATGAGATTTCTT 960  
QY 961 TCAATTTTTTACTGCAAGTTTATTCGACGATCCTCCGCAATTAGCTGCTCCAGTCAACT 1020  
DB 961 TCAATTTTTTACTGCAAGTTTATTCGACGATCCTCCGCAATTAGCTGCTCCAGTCAACT 1020  
QY 1021 AGACAGAGATGAACCGCACAAATTCGGCTGAGCTGATCGTTCATCGTTTACTCAGATT 1080  
DB 1021 AGACAGAGATGAACCGCACAAATTCGGCTGAGCTGATCGTTCATCGTTTACTCAGATT 1080  
QY 1081 GAAGGGGATTTTCGATGTTGCTGTTTGGCAATTTTCCAAACAGCACAAATAACGGGTTAT 1140  
DB 1081 GAAGGGGATTTTCGATGTTGCTGTTTGGCAATTTTCCAAACAGCACAAATAACGGGTTAT 1140  
QY 1141 TTTTAAATACTACTATTTGCCAGCATTTGCTGCTAAAGAGAGGGTATCTCTCAGAAA 1200  
DB 1141 TTTTAAATACTACTATTTGCCAGCATTTGCTGCTAAAGAGAGGGTATCTCTCAGAAA 1200  
QY 1201 AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATG 1260  
DB 1201 AGAGAGGCTGAAGCTTA----- 1217  
QY 1261 GCGCCCGGAGAGGATATAATTTGGGCTTCCCGGCTGTCACGCCAGCAGTGCAGGAG 1320  
DB 1218 ----- 1217  
QY 1321 AGAGGTTGCTGTTTGAATGACAGTGTCCGGGATTTCCCGTGTCTTCCACCCCATGGCC 1380  
DB 1218 ----- 1217  
QY 1381 ATCGAACAACCTCAAGAGAGAGATGTCCTTTCTAACTAGTGGCGTAGAATTTCCCTAGGG 1440  
DB 1218 -----CGTAGAATTTCCCTAGGG 1234  
QY 1441 CGCCCGGAAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTTGGGCACTTAGC 1500  
DB 1235 CGCCCGGAAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTTGGGCACTTAGC 1294  
QY 1501 AGAAGACCCGCTCTGTAGATTTCTAATCAAGAGGATGTCAAGATGCCATTTGCCTGAGAG 1560  
DB 1295 AGAAGACCCGCTCTGTAGATTTCTAATCAAGAGGATGTCAAGATGCCATTTGCCTGAGAG 1354  
QY 1561 ATCGAGCTTCATTTTGTATACCTTTTATTTTAACTTAACTATATAGTATAGGATTTTTTT 1620  
DB 1355 ATCGAGCTTCATTTTGTATACCTTTTATTTTAACTTAACTATATAGTATAGGATTTTTTT 1414  
QY 1621 GTCATTTGTTTCTCTGCTAGAGCTTCTCTGATCAGCTATCTCCAGCTGATGAA 1680  
DB 1415 GTCATTTGTTTCTCTGCTAGAGCTTCTCTGATCAGCTTCTCCAGCTGATGAA 1474  
QY 1681 TATCTTTGCTAGGGGTTTGGGAAATCAATTCAGTTTGCATGTTTCTTGTATTTTCC 1740  
DB 1475 TATCTTTGCTAGGGGTTTGGGAAATCAATTCAGTTTGCATGTTTCTTGTATTTTCC 1534  
QY 1741 ACTCCTTTTCAGAGTACAGAGATTAAGTGAGAGTTGCTTTGTGCAAGCTTATCGATAA 1800  
DB 1535 ACTCCTTTTCAGAGTACAGAGATTAAGTGAGAGTTGCTTTGTGCAAGCTTATCGATAA 1594  
QY 1801 GCTTTAATGCGGTAGTTTATCAGTTAAATTTGCTAAACGAGTTCAGGCAACCGTGTATGAA 1860  
DB 1595 GCTTTAATGCGGTAGTTTATCAGTTAAATTTGCTAAACGAGTTCAGGCAACCGTGTATGAA 1654  
QY 1861 ATCTAAATGCGGCTCATGCTATCTCCGCAACCGTCACTCCGATGCTGTAGGCAATAG 1920  
DB 1655 ATCTAAATGCGGCTCATGCTATCTCCGCAACCGTCACTCCGATGCTGTAGGCAATAG 1714  
QY 1921 CTTGGTTATGCGGTACTCGCGGCTCTTGGGGATATCGTCCATTCGACAGATCGC 1980  
DB 1715 CTTGGTTATGCGGTACTCGCGGCTCTTGGGGATATCGTCCATTCGACAGATCGC 1774  
QY 1981 CAGTCACTATGCGGCTCTCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCT 2040  
DB 1775 CAGTCACTATGCGGCTCTCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCT 1834

QY 2041 TCTCGGAGCACTGTCGACCGCTTTGGCGCGCCGAGTCTGCTCGCTTCGCTACTTGG 2100  
Db 1835 TCTCGGAGCACTGTCGACCGCTTTGGCGCGCCGAGTCTGCTCGCTTCGCTACTTGG 1894  
QY 2101 AGGCACATATCGACTACGCGATCATGGCGACCAACCCGTCCTGTGATCATCGAATCTA 2160  
Db 1895 AGGCACATATCGACTACGCGATCATGGCGACCAACCCGTCCTGTGATCATCGAATCTA 1954  
QY 2161 AATGTAAAGTAAATCTCTAAATTAATTAATAGTCCAGTTCTCCATACGAACTTAA 2220  
Db 1955 AATGTAAAGTAAATCTCTAAATTAATTAATAGTCCAGTTCTCCATACGAACTTAA 2014  
QY 2221 CAGCAATTCGGGTAGCACTCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2280  
Db 2015 CAGCAATTCGGGTAGCACTCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2074  
QY 2281 TATGTTTCAGTCCCTCAGGAGTTACGTTCTGTGAAGTGAATCTTCTGAAGTTGCGAG 2340  
Db 2075 TATGTTTCAGTCCCTCAGGAGTTACGTTCTGTGAAGTGAATCTTCTGAAGTTGCGAG 2134  
QY 2341 TGTAACTCCGCTGATTTGACGGGCATATCCGTACGTTGGCAAGTGTGGTTCGTAACCGG 2400  
Db 2135 TGTAACTCCGCTGATTTGACGGGCATATCCGTACGTTGGCAAGTGTGGTTCGTAACCGG 2194  
QY 2401 AGGAGTAATCTCCAACCTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTGT 2460  
Db 2195 AGGAGTAATCTCCAACCTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTGT 2254  
QY 2461 GTACTGATCAACATAGAGAGCAATCTCGATTTGCAGGATCAAGTGTTCAGGAGCGT 2520  
Db 2255 GTACTGATCAACATAGAGAGCAATCTCGATTTGCAGGATCAAGTGTTCAGGAGCGT 2314  
QY 2521 ACTGATTTGGACATTTCCAAAGCCCTGCTGAGGTTGCAACCGATAGGGTTGTAGAGTGT 2580  
Db 2315 ACTGATTTGGACATTTCCAAAGCCCTGCTGAGGTTGCAACCGATAGGGTTGTAGAGTGT 2374  
QY 2581 CAATACACTTCGGTACATTTCAACCCCTTGGCACTGCAGCTGCTGTTGTAACAGCAT 2640  
Db 2375 CAATACACTTCGGTACATTTCAACCCCTTGGCACTGCAGCTGCTGTTGTAACAGCAT 2434  
QY 2641 CTTCATTTCTGGCAAGTCTCTGCTCATATCGACAGCAACAGCAATCACCTGGGAAT 2700  
Db 2435 CTTCATTTCTGGCAAGTCTCTGCTCATATCGACAGCAACAGCAATCACCTGGGAAT 2494  
QY 2701 CAATACCATTTAGCTTGAGACAGAGGCTTGAGGCAACGAAATCTGGATCAGCGTATT 2760  
Db 2495 CAATACCATTTAGCTTGAGACAGAGGCTTGAGGCAACGAAATCTGGATCAGCGTATT 2554  
QY 2761 TATCAGCAATTAACATAGCACTTCAGAGGCCAGCAGCATGTCAATACTACACAGGCGTG 2820  
Db 2555 TATCAGCAATTAACATAGCACTTCAGAGGCCAGCAGCATGTCAATACTACACAGGCGTG 2614  
QY 2821 ATGTGTCAATTTGAACCATCATCTTGGCAGCAGTAAAGAACTGGTTTCCTGGCCAAATA 2880  
Db 2615 ATGTGTCAATTTGAACCATCATCTTGGCAGCAGTAAAGAACTGGTTTCCTGGCCAAATA 2674  
QY 2881 TTTTGTCACTTAGGACAGTTCTGTTCCGTAAGCCATAGCAGTACTGCTGGGCGC 2940  
Db 2675 TTTTGTCACTTAGGACAGTTCTGTTCCGTAAGCCATAGCAGTACTGCTGGGCGC 2734  
QY 2941 CTCCTGCTAGCAGCATACATTTAGCAACCACTTTGTGGGCAACGTAGATCACTTCCTGGG 3000  
Db 2735 CTCCTGCTAGCAGCATACATTTAGCAACCACTTTGTGGGCAACGTAGATCACTTCCTGGG 2794  
QY 3001 TAAAGGTACCATCTCTTCTAGGTGAGATGCAAAAACAAATTTCTTTCGACACGCACTT 3060  
Db 2795 TAAAGGTACCATCTCTTCTAGGTGAGATGCAAAAACAAATTTCTTTCGACACGCACTT 2854  
QY 3061 TGGCAGGAACCCAGCATCAGGGAAGTGAAGGAGCAAAATTCGGTTCCACAGGAATAT 3120  
Db 2855 TGGCAGGAACCCAGCATCAGGGAAGTGAAGGAGCAAAATTCGGTTCCACAGGAATAT 2914

QY 3121 AGAGCCCAACTTTCTCAATAGGTCTTGCAAAACGAGAGCAGACTACACAGGGGCAAGTCT 3180  
Db 2915 AGAGCCCAACTTTCTCAATAGGTCTTGCAAAACGAGAGCAGACTACACAGGGGCAAGTCT 2974  
QY 3181 CAATCTGCAACGCTCTCGTTAGTTAGCTTCATGGAATTTCTTGACGCTTATCTATAGAGA 3240  
Db 2975 CAATCTGCAACGCTCTCGTTAGTTAGCTTCATGGAATTTCTTGACGCTTATCTATAGAGA 3034  
QY 3241 GATCAATGGCTCTCTTAAACGTTATCTGGCAATTCATPAAGTTCCTCTGGAAAGGAGCTT 3300  
Db 3035 GATCAATGGCTCTCTTAAACGTTATCTGGCAATTCATPAAGTTCCTCTGGAAAGGAGCTT 3094  
QY 3301 CTAAACAGGTGCTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTT 3360  
Db 3095 CTAAACAGGTGCTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTT 3154  
QY 3361 TGTCAACATTTTGACGAACATTTGTCGAATTTGACTTAATTCATATCTGTTCCG 3420  
Db 3155 TGTCAACATTTTGACGAACATTTGTCGAATTTGACTTAATTCATATCTGTTCCG 3214  
QY 3421 TTTTCTGATAGGACGACGAAGGGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAGCT 3480  
Db 3215 TTTTCTGATAGGACGACGAAGGGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAGCT 3274  
QY 3481 CAATTTTGCACAAATTCATACGACCTTCCAGAGGCACTTCTTTAGGTTTGGATTTCTT 3540  
Db 3275 CAATTTTGCACAAATTCATACGACCTTCCAGAGGCACTTCTTTAGGTTTGGATTTCTT 3334  
QY 3541 TAGGTTGTTCTTGGTGATCTCTGGTTGGCATCTCCTTTCTCTAGTACCTTTAGG 3600  
Db 3335 TAGGTTGTTCTTGGTGATCTCTGGTTGGCATCTCCTTTCTCTAGTACCTTTAGG 3394  
QY 3601 ACTTCATATCAGGTTCTCTCCACCTCGTCCACGTCACACCGTACTTGCACATCTAA 3660  
Db 3395 ACTTCATATCAGGTTCTCTCCACCTCGTCCACGTCACACCGTACTTGCACATCTAA 3454  
QY 3661 CTAAATGCAAAATAAAATAAGTCAGCACATTTCCAGGCTATATCTTCTGATTTAGCTT 3720  
Db 3455 CTAAATGCAAAATAAAATAAGTCAGCACATTTCCAGGCTATATCTTCTGATTTAGCTT 3514  
QY 3721 CTGGAAGTTCAATCAGCTTCTTCCCTTAATTTTGGGTTTCAACAAACTTCGTCGTCATA 3780  
Db 3515 CTGGAAGTTCAATCAGCTTCTTCCCTTAATTTTGGGTTTCAACAAACTTCGTCGTCATA 3574  
QY 3781 ACCGTTTGGTATAGAACCTTCTGGAGCATTTGCTTACGATCCCAAGGTTGCTTCCA 3840  
Db 3575 ACCGTTTGGTATAGAACCTTCTGGAGCATTTGCTTACGATCCCAAGGTTGCTTCCA 3634  
QY 3841 TGGCTCTAAGACCTTTGATTTGGCCAAACAGGAGTGGCTTCCAAGTGACAGAAACCA 3900  
Db 3635 TGGCTCTAAGACCTTTGATTTGGCCAAACAGGAGTGGCTTCCAAGTGACAGAAACCA 3694  
QY 3901 CACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCAATCCAAATTCGATACCCA 3960  
Db 3695 CACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCAATCCAAATTCGATACCCA 3754  
QY 3961 GCAACTTTTGTGCTCCAGATGAGCACCTTTATACCAAAACCGTGACGAGAT 4020  
Db 3755 GCAACTTTTGTGCTCCAGATGAGCACCTTTATACCAAAACCGTGACGAGAT 3814  
QY 4021 GGTAGACTCAGTTTGTCTCTTATAGCTCCGGAATAGACTTTTGGCAGGATACCA 4080  
Db 3815 GGTAGACTCAGTTTGTCTCTTATAGCTCCGGAATAGACTTTTGGCAGGATACCA 3874  
QY 4081 GGCCCAACGAGTAATTAGAAGAGTCAGCCACCAAGTAGTGAATAGACCATTCGGGGCGT 4140  
Db 3875 GGCCCAACGAGTAATTAGAAGAGTCAGCCACCAAGTAGTGAATAGACCATTCGGGGCGT 3934  
QY 4141 CAGTAGTCAAGACCCCAACAAATTTTCACTGACAGGAACTTTTGTGACATCTTCAGAAA 4200  
Db 3935 CAGTAGTCAAGACCCCAACAAATTTTCACTGACAGGAACTTTTGTGACATCTTCAGAAA 3994  
QY 4201 GTTCGTATTTAGTAGTCAATTTGCCGAGCATCAATTAATGGGATTTATACCGAAGCAACAG 4260

Db 3995 GTTCGTTAGTAGTCAATTCGCGAGCATCAATATGGGGATTATACGAGCAACAG 4054  
Qy 4261 TGGAAATGACATCTACCAACTTTTCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4320  
Db 4055 TGGAAATGACATCTACCAACTTTTCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4114  
Qy 4321 CATTAAGTGAATCTTTTCAAAATCGCCAGTGGAGAGAAAAGGACACAGCGATATAGCAT 4380  
Db 4115 CATTAAGTGAATCTTTTCAAAATCGCCAGTGGAGAGAAAAGGACACAGCGATATAGCAT 4174  
Qy 4381 TAGCGGCAAGATGCAACTTTTATCAACAGGGTCTTATAGATAACCCCTAGCGGCTGGGA 4440  
Db 4175 TAGCGGCAAGATGCAACTTTTATCAACAGGGTCTTATAGATAACCCCTAGCGGCTGGGA 4234  
Qy 4441 TCATCTTTGACAACTCTTTCTGCCAAATCTAGTCCCAAAATCAATCAATCTTATGATACCAT 4500  
Db 4235 TCATCTTTGACAACTCTTTCTGCCAAATCTAGTCCCAAAATCAATCAATCTTATGATACCAT 4294  
Qy 4501 TATTGTCAAACTTGAGCAAGTTGTGATCAGCTCCTCAAAATTTGCTCTGTACCGGATG 4560  
Db 4295 TATTGTCAAACTTGAGCAAGTTGTGATCAGCTCCTCAAAATTTGCTCTGTACCGGATG 4354  
Qy 4561 ACTCAACTTGACAACTTTTAAAGCTCAGTCTGATGAGTGAATCTTATCAGTGTGTGCA 4620  
Db 4355 ACTCAACTTGACAACTTTTAAAGCTCAGTCTGATGAGTGAATCTTATCAGTGTGTGCA 4414  
Qy 4621 GCTGTCAGCAGCATAGGAAACACAGGCTTTCTCCTACCAACTCAAGGAATATCAAACT 4680  
Db 4415 GCTGTCAGCAGCATAGGAAACACAGGCTTTCTCCTACCAACTCAAGGAATATCAAACT 4474  
Qy 4681 CTGCAACTTTGCGTATGACAGTGAAGAAATGTCATCTTGAAGTGGACAGTGAG 4740  
Db 4475 CTGCAACTTTGCGTATGACAGTGAAGAAATGTCATCTTGAAGTGGACAGTGAG 4534  
Qy 4741 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATATCAGTGAATCTCAGTGTATCAGGAGA 4800  
Db 4535 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATATCAGTGAATCTCAGTGTATCAGGAGA 4594  
Qy 4801 TCCTCTACGCGGAGCATCTGTGCGCA 4828  
Db 4595 TCCTCTACGCGGAGCATCTGTGCGCACCTGAGGGGGGGGGGGCGGTGAGTCTGC 4654  
Qy 4829 4828  
Db 4655 CTCGTGAAGAAGTGTGCTGACTCATACAGGCGCTGAATCGCCCCATCATCCAGCAGA 4714  
Qy 4829 4828  
Db 4715 AAGTGAAGGACCCAGGTTGATGAGAGCTTTGTTCTAGGTGACAGCTGTTGATTTGA 4774  
Qy 4829 4828  
Db 4775 ACTTTTGTCTTGGCACGGAACGGTCTGGGTTGTGCGGAAGATGGGTGATCTGATCCTTCA 4834  
Qy 4829 4828  
Db 4835 ACTGACAAAGTTCGATTTATTAACAAAGCCGCGCTCCGTCAGTCAAGTCAAGTATGCT 4894  
Qy 4829 4828  
Db 4895 CTGCCAGTGTATACCAATTAACCAATTCGATTAGAAAAAATCATCGAGCATCAAAATG 4954  
Qy 4829 4828  
Db 4955 AAATGCAATTTATTATATCATACAGGATATATCAATACCATATTTTGA AAAAGCCGTTTCTG 5014  
Qy 4829 4828  
Db 5015 TAAATGAGGAGAAACTCACCAGGCGAGTTCATAGGATGGCAAGATCCTGGTATCGGTC 5074  
Qy 4829 4828

5075 TCGAATTCGACTCGTCCAAACATCAATACAACTTATTAATTTCCCTCTGCTCAAAATAAG 5134  
4829 4828  
5135 GTTATCAAGTGAATAATCACCATGAGTGAAGTGAATCCGCTGAGATGGCAAAAGCTT 5194  
4829 4828  
5195 ATGCATTTCTTTCCAGACTTTGTTCAACAGGCGAGCATTTAGCTGCTCATCAAAATCACT 5254  
4829 4828  
5255 CGCATCAACCAAAACCGTTTATTCATTCTGTGATTGCGCTGAGGAGAGCAAAATACGGGATC 5314  
4829 4828  
5315 GCTGTTAAAGGACAAATTACAAACAGGAATCGAATGCAACCGGCGAGCAACACTGCCAG 5374  
4829 4828  
5375 CGCATCAACAAATATTTTACCTGAATCAGGATATTCTTTCTAATACCTGGAATGCTGTTTT 5434  
4829 4828  
5435 CCGGGGATCGCAGTGGTGAATACCATGCATCATCAGGATACGATATAAATGCTTGAT 5494  
4829 4828  
5495 GGTGGAAGAGGCAATAATTCGTCAGCAGTTTAGTCTGACCACTCTCATCTGTAACATC 5554  
4829 4828  
5555 ATTGGAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGCGGCAATCGGGCTTCCCAT 5614  
4829 4828  
5615 CAATCGATAGATTGTGCGACCTGATTGCGCGACATTATCGGAGCCCATTTATACCCATA 5674  
4829 4828  
5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGCTCGAGCAAGAGCTTCCCGTTGAAT 5734  
4829 4828  
5735 ATGGTCAATAACCCCTTGTATTACTGTTTATGTATGTAAGCAGACAGTTTATTGTTTCATGA 5794  
4829 4828  
5795 TGATATATTTTATCTTTGCAATGTAAATCAGAGATTTTGAGACACAACTGGGCTTTC 5854  
4829 4828  
5855 CCCCCCCCCCTGTCAGTGGGAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTGTTT 5914  
4880 CGCGACATCACCGATGGGAAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTGTTT 4939  
5915 CGCGACATCACCGATGGGAAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTGTTT 5974  
4940 CGCGCTGGGTATGTTGGAGGCGGCTGCGCGGAGACTGTTGGGCGGCTCATCTCTTGA 4999  
5975 CGCGCTGGGTATGTTGGAGGCGGCTGCGCGGAGACTGTTGGGCGGCTCATCTCTTGA 6034  
5000 TGCAACATTTCTTGGCGGCGGCTGCTCAACGCGCTCAACCTACTACTGCGGCTGCTTCT 5059  
6035 TGCAACATTTCTTGGCGGCGGCTGCTCAACGCGCTCAACCTACTACTGCGGCTGCTTCT 6094  
5060 AATGCAAGGATCGCATTAAGGAGAGCGTGAATATCTATGATTGGAAGTATGGGAATGGT 5119  
6095 AATGCAAGGATCGCATTAAGGAGAGCGTGAATATCTATGATTGGAAGTATGGGAATGGT 6154  
5120 GATACCGGATCTTCTCAGTGTCTTGAAGTCTCTATCAGATTATGCCAACTAAAGCAAC 5179  
6155 GATACCGGATCTTCTCAGTGTCTTGAAGTCTCTATCAGATTATGCCAACTAAAGCAAC 6214



Db 8375 TAAGTAGTTCGCCAGTTAATAGTTGGCAACGGTTGTGCAATTCGTCAGGCAATCGTGG 8434  
 Qy 7400 TGTACCGCTCGCTGGTTGGTATGGCTTCATTACAGTCCGGTCCCAACGATCAAGGCGAG 7459  
 Db 8435 TGTACCGCTCGCTGGTATGGCTTCATTACAGTCCGGTCCCAACGATCAAGGCGAG 8494  
 Qy 7460 TTAATGATCCCAATGTTGTGCAAAAGCGGTTAGTCTCGTCCCTCCGATCGTTG 7519  
 Db 8495 TTACATGATCCCAATGTTGTGCAAAAGCGGTTAGTCTCGTCCCTCCGATCGTTG 8554  
 Qy 7520 TCAGAAGTAAGTGGCCGACGTGTTATCACTCATGTTATGGCAGCTGCATATTTCTC 7579  
 Db 8555 TCAGAAGTAAGTGGCCGACGTGTTATCACTCATGTTATGGCAGCTGCATATTTCTC 8614  
 Qy 7580 TTACTGTCATGCCATCCCTAGATGCTTTTCTGTGACGTGTTAGTCTCAACCAAGTCAT 7639  
 Db 8615 TTACTGTCATGCCATCCCTAGATGCTTTTCTGTGACGTGTTAGTCTCAACCAAGTCAT 8674  
 Qy 7640 TCTGAGATAGTATGCGGACCGAGTTCCTTTCGCCGCGCTCAACCGGATATA 7699  
 Db 8675 TCTGAGATAGTATGCGGACCGAGTTCCTTTCGCCGCGCTCAACCGGATATA 8734  
 Qy 7700 CGCGCCCATAGCAGAACTTTAAAGTGTCTCATCATTTGAAAGCTTTCTTCGGGGCGAA 7759  
 Db 8735 CGCGCCCATAGCAGAACTTTAAAGTGTCTCATCATTTGAAAGCTTTCTTCGGGGCGAA 8794  
 Qy 7760 AACTCTCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTGTCACCCA 7819  
 Db 8795 AACTCTCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTGTCACCCA 8854  
 Qy 7820 ACTGATCTTACGATCTTTACTTTTACAGAGCTTTCTGGGTGAGCAAAAACAGGAAGGC 7879  
 Db 8855 ACTGATCTTACGATCTTTACTTTTACAGAGCTTTCTGGGTGAGCAAAAACAGGAAGGC 8914  
 Qy 7880 AAATGCGCGCAAAAAGGGAATAAGGCGCACCGGAATGTTGAATCTATCTATCTTCC 7939  
 Db 8915 AAATGCGCGCAAAAAGGGAATAAGGCGCACCGGAATGTTGAATCTATCTATCTTCC 8974  
 Qy 7940 TTTTTCATATTTATGAGCATTTATCATGAGGTTATTTCTCATGAGCGGATACATATTTG 7999  
 Db 8975 TTTTTCATATTTATGAGCATTTATCATGAGGTTATTTCTCATGAGCGGATACATATTTG 9034  
 Qy 8000 AATGATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTGCCAC 8059  
 Db 9035 AATGATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTGCCAC 9094  
 Qy 8060 CTGAGCTCTAAGAACCATTTATTCATGACATTAACCTATAAATAAGGCGTATCAGA 8119  
 Db 9095 CTGAGCTCTAAGAACCATTTATTCATGACATTAACCTATAAATAAGGCGTATCAGA 9154  
 Qy 8120 GGCCCTTTCGTTCTCAAGAAATTAATTTCTCATGTTTGACAGCTTTATCATCGATAAGCTGAC 8179  
 Db 9155 GGCCCTTTCGTTCTCAAGAAATTAATTTCTCATGTTTGACAGCTTTATCATCGATAAGCTGAC 9214  
 Qy 8180 TCATGTTGTTATGTAATAAGCAGATCGGGAACCTGAAATAATACAGTTATTTATT 8239  
 Db 9215 TCATGTTGTTATGTAATAAGCAGATCGGGAACCTGAAATAATACAGTTATTTATT 9274  
 Qy 8240 CG 8241  
 Db 9275 CG 9276

RESULT 9  
 AB222459  
 ID AB222459 standard; DNA; 9276 BP.  
 XX AB222459;  
 AC  
 XX  
 XX  
 XX 24-MAR-2003 (first entry)  
 XX Pichia pastoris vector pPIC9K comprising MTSP10 DNA SEQ ID NO:13.  
 DE

XX Human; membrane-type serine protease; enzyme: MTSP10; cytosolic;  
 KW type-II membrane-type serine protease; neoplastic disease; tumour; gene;  
 KW da.  
 XX Pichia pastoris.  
 OS Homo sapiens.  
 OS Synthetic.  
 XX WO200292841-A2.  
 PN 21-NOV-2002.  
 XX PD 14-MAY-2002; 2002WO-US015332.  
 XX PF 14-MAY-2001; 2001US-0291001P.  
 PR (CORV-) CORVAS INT INC.  
 XX Madison EL, Yeh J;  
 XX WPI; 2003-129309/12.  
 XX New polypeptides comprising the protease domain of a type-II membrane-  
 PT type serine protease (MTSP10), or its muteins, useful for diagnosing  
 PT neoplasms or malignancies, or for screening for MTSP10 inhibitors for  
 PT treating such diseases.  
 XX Example 2; Page 189-191; 198pp; English.  
 PS The present invention describes a polypeptide comprising a purified  
 CC single or two chain polypeptide, which comprises the protease domain of a  
 CC type-II membrane-type serine protease (MTSP10) or its catalytically  
 CC active portion, or a mutain of it, where up to 50 % of the amino acids  
 CC are replaced with another amino acid, and the resulting polypeptide is a  
 CC single chain or two chain polypeptide that has a catalytic activity of at  
 CC least 1-10 % of the unmutated polypeptide. MTSP10 has cytostatic  
 CC activity. The polypeptide containing the protease domain of the MTSP10 is  
 CC useful for detecting a neoplastic disease, and for diagnosing the  
 CC presence of a pre-malignant lesion, a malignancy, or other pathologic  
 CC condition in a subject, or monitoring tumour (e.g. breast, cervix,  
 CC prostate, lung, ovary or colon tumour) progression and/or therapeutic  
 CC effectiveness. An inhibitor of the polypeptide containing the protease  
 CC domain of MTSP10 is useful for treating or preventing neoplastic disease  
 CC in a mammal. An inhibitor of the activation cleavage of the zymogen form  
 CC of the MTSP10 polypeptide is useful for inhibiting tumour initiation,  
 CC growth or progression, or treating (pre-)malignant conditions of the e.g.  
 CC breast, cervix, prostate, lung, ovary or colon. The present sequence  
 CC represents a Pichia pastoris vector designated pPIC9K, comprising MTSP10,  
 CC which is used in an example from the present invention  
 XX Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;

Query Match 79.6%; Score 6561.6; DB 7; Length 9276;  
 Best Local Similarity 84.7%; Pred. No. 0;  
 Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;  
 Qy 1 AGATCTAACATCCAAAGACGAAAGGTTGATGAACCTTTTCCCATCCGACATCCACAG 60  
 Db 1 AGATCTAACATCCAAAGACGAAAGGTTGATGAACCTTTTCCCATCCGACATCCACAG 60  
 Qy 61 GTCCATTTCTCACATAGTGCACAAACGCAACAGAGGGGATACACTAGCAGACAGCGT 120  
 Db 61 GTCCATTTCTCACATAGTGCACAAACGCAACAGAGGGGATACACTAGCAGACAGCGT 120  
 Qy 121 TGCAAAACGAGGACCTCCACT 180  
 Db 121 TGCAAAACGAGGACCTCCACT 180  
 Qy 181 AGCCAGTTATGGGCTTGATGGAGCTCGCTCATTCCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
 Db 181 AGCCAGTTATGGGCTTGATGGAGCTCGCTCATTCCTCTCTCTCTCTCTCTCTCTCTCTCT 240

QY 241 ACACATGACCTTTATAGCCGTCTATCCCTGGCCCCCTCGCGAGGTTTCATGTTTGT 300  
DB 241 ACACATGACCTTTATAGCCGTCTATCCCTGGCCCCCTCGCGAGGTTTCATGTTTGT 300  
QY 301 TTTCCGAATGCAACAAGCTCCGCAATTAACACCGCAACATCACTCCAGATGAGGCTTTCTG 360  
DB 301 TTTCCGAATGCAACAAGCTCCGCAATTAACACCGCAACATCACTCCAGATGAGGCTTTCTG 360  
QY 361 AGTGTGGGTCAAAATAGTTTCATGTTCCCAATGGCCCAAACTGACGTTTAAAGCCT 420  
DB 361 AGTGTGGGTCAAAATAGTTTCATGTTCCCAATGGCCCAAACTGACGTTTAAAGCCT 420  
QY 421 GTCCTTGGAACTTAATATGACAAAAGCGTGATCTCATCAAGATGAACCTAAGTTTGGTTTCG 480  
DB 421 GTCCTTGGAACTTAATATGACAAAAGCGTGATCTCATCAAGATGAACCTAAGTTTGGTTTCG 480  
QY 481 TTGAAATGCTAAACGCGCAGTTGGTCAAAAAGAACTTCCAAAAGTCGCCATACCGTTTGT 540  
DB 481 TTGAAATGCTAAACGCGCAGTTGGTCAAAAAGAACTTCCAAAAGTCGCCATACCGTTTGT 540  
QY 541 CTGTTTGGTATGATGACGAATGCTCAAAAATAAATCTCATTAATGCTTAGCGCAGTCT 600  
DB 541 CTGTTTGGTATGATGACGAATGCTCAAAAATAAATCTCATTAATGCTTAGCGCAGTCT 600  
QY 601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGAAACGCAATGGGAAACACCCGCT 660  
DB 601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGAAACGCAATGGGAAACACCCGCT 660  
QY 661 TTTTGGATGATGATGCTCCCAATGCTATGCTTCCCAAGATTCCTGGTGGGAATACT 720  
DB 661 TTTTGGATGATGATGCTCCCAATGCTATGCTTCCCAAGATTCCTGGTGGGAATACT 720  
QY 721 GCTGATAGCTAAAGCTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780  
DB 721 GCTGATAGCTAAAGCTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780  
QY 781 ATATAAACAGAGAAGCTGCTGTTTAAACCTTTTATATCATCATTAATAGCTT 840  
DB 781 ATATAAACAGAGAAGCTGCTGTTTAAACCTTTTATATCATCATTAATAGCTT 840  
QY 841 ACTTTCAATTTGGACCTGTTTCAATGACAACTTTGATTTTAAACGACTTTTAAACGA 900  
DB 841 ACTTTCAATTTGGACCTGTTTCAATGACAACTTTTGAATTTAAACGACTTTTAAACGA 900  
QY 901 CAACCTTGAGAAGATCAAAAAACAATTAATTCGAAGGATCCAAACGATGAGATTTCCCT 960  
DB 901 CAACCTTGAGAAGATCAAAAAACAATTAATTCGAAGGATCCAAACGATGAGATTTCCCT 960  
QY 961 TCAATTTTACTGCAAGTTTATTCGACGATCTCCGCAATTAGCTGCTCCAGTCAACACT 1020  
DB 961 TCAATTTTACTGCAAGTTTATTCGACGATCTCCGCAATTAGCTGCTCCAGTCAACACT 1020  
QY 1021 ACAACAGAAGATGAACCGGCAAAATTCGCGTGAAGCTGTCATPCGTTTACTCAGATTTA 1080  
DB 1021 ACAACAGAAGATGAACCGGCAAAATTCGCGTGAAGCTGTCATPCGTTTACTCAGATTTA 1080  
QY 1081 GAAGGGATTTGATGTTGCTGTTTGGCAATTTCCACACGACCAATTAAGGGTTATG 1140  
DB 1081 GAAGGGATTTGATGTTGCTGTTTGGCAATTTCCACACGACCAATTAAGGGTTATG 1140  
QY 1141 TTTTATAATATCTATTTGCAAGATGCTGCTGCTTAAAGAAGGGGTATCTCTCGAGAAA 1200  
DB 1141 TTTTATAATATCTATTTGCAAGATGCTGCTGCTTAAAGAAGGGGTATCTCTCGAGAAA 1200  
QY 1201 AGAGAGGCTGAAGCCCGAGCCAGCCAGCCAGCCAGCCAGGAAACAATGATCATG 1260  
DB 1201 AGAGAGGCTGAAGCTTA----- 1217  
QY 1261 GCCCCCCGGGAGAGGATAAATTTGGCTTCCCGGTGTCCCGGCTGACCGCCAGCAGTGCACGGAG 1320  
DB 1218 ----- 1217  
QY 1321 AGAGGTTCTGTTTGTATGACAGTGTCCGGGATTCCTGGTGGTCTCCGCCCATCGCC 1380

DB 1218 ----- 1217  
QY 1381 ATCGAACAACACTCAAGAAGAATGTCCCTTTCTAACTAGTGGGTAGAAATTCCTAGGG 1440  
DB 1218 -----CGTAGAATTCCTAGGG 1234  
QY 1441 CGCGCGCAATTAATTCGCTTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTTACG 1500  
DB 1235 CGCGCGCAATTAATTCGCTTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTTACG 1294  
QY 1501 AGAAGACCGGCTTCTGCTAGATCTTAATCAAGAGATGTCAGAAATGCCATTTGCCGTGAGAG 1560  
DB 1295 AGAAGACCGGCTTCTGCTAGATCTTAATCAAGAGATGTCAGAAATGCCATTTGCCGTGAGAG 1354  
QY 1561 ATGCAAGGCTTCATTTTTCATATCTTTTATTTGTAACCTATATAGTATAGGATTTT 1620  
DB 1355 ATGCAAGGCTTCATTTTTCATATCTTTTATTTGTAACCTATATAGTATAGGATTTT 1414  
QY 1621 GTCAATTTGTTTCTCTCGTACGAGCTTCTCTCATCAGCCTATCTCGCAGCTGATGAA 1680  
DB 1415 GTCAATTTGTTTCTCTCGTACGAGCTTCTCTCATCAGCCTATCTCGCAGCTGATGAA 1474  
QY 1681 TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGTGTTTCTTCTGATTTTCCC 1740  
DB 1475 TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGTGTTTCTTCTGATTTCCC 1534  
QY 1741 ACTCTCTCTCAGAGTACAGAAATTAAGTGAGAAAGTTTGTGCAAGCTTATCGATAA 1800  
DB 1535 ACTCTCTCTCAGAGTACAGAAATTAAGTGAGAAAGTTTGTGCAAGCTTATCGATAA 1594  
QY 1801 GCTTAAATCGGCTAGTTTATCAGATTAATTTGTAACGAGTCAAGCACCGTGTATGAA 1860  
DB 1595 GCTTAAATCGGCTAGTTTATCAGATTAATTTGTAACGAGTCAAGCACCGTGTATGAA 1654  
QY 1861 ATCTAAACATGGCTCATCTCTCGCACCGTCAACCTGGATGCTGTAGGCATAGG 1920  
DB 1655 ATCTAAACATGGCTCATCTCTCGCACCGTCAACCTGGATGCTGTAGGCATAGG 1714  
QY 1921 CTTGTTTATGCGGTACTGCGGGGCTCTTGGGGATATCTGTCATTCGACAGATCGC 1980  
DB 1715 CTTGTTTATGCGGTACTGCGGGGCTCTTGGGGATATCTGTCATTCGACAGATCGC 1774  
QY 1981 CAGTCACTATGCGGTCTGCTAGCGTATATGCGTTGATGCAATTTCTATCGCACCGCT 2040  
DB 1775 CAGTCACTATGCGGTCTGCTAGCGTATATGCGTTGATGCAATTTCTATCGCACCGCT 1834  
QY 2041 TCTCGAGCAGTGTCCGACCGCTTTGGCGCGCCCGCTCTGCTCTGCTTCTGCTACTTGG 2100  
DB 1835 TCTCGAGCAGTGTCCGACCGCTTTGGCGCGCCCGCTCTGCTCTGCTTCTGCTACTTGG 1894  
QY 2101 AGCCATATCGACTAGCGATCATGGGACACACACCGCTGCTGATCTCGAATCTA 2160  
DB 1895 AGCCATATCGACTAGCGATCATGGGACACACACCGCTGCTGATCTCGAATCTA 1954  
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2220  
DB 1955 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2014  
QY 2221 CAGCATTTGGGTGAGCATCTAGACCTTCAAACAGCGCCAGATCCATCTGTTGGCCAA 2280  
DB 2015 CAGCATTTGGGTGAGCATCTAGACCTTCAAACAGCGCCAGATCCATCTGTTGGCCAA 2074  
QY 2281 TATGTTTTCAGTCCCTCAGAGTTAGCTCTTGTGAAGTGAATCTCTGGAAGGTTGCGAG 2340  
DB 2075 TATGTTTTCAGTCCCTCAGAGTTAGCTCTTGTGAAGTGAATCTCTGGAAGGTTGCGAG 2134  
QY 2341 TGTAACTCCGCTGTATTTGACGGGATATCCGTAGCTTTGGCAAAAGTGTGTTGATACGG 2400  
DB 2135 TGTAACTCCGCTGTATTTGACGGGATATCCGTAGCTTTGGCAAAAGTGTGTTGATACGG 2194  
QY 2401 AGGAGTAACTCCCAACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGCGTGT 2460



Db 2195 AGGAGTAATCCCAACTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTGT 2254  
Qy 2461 GTACTGATCAACATAGAGAGCAATCTCGATTTGCAGATCAAGTGTTTCAGGAGCGT 2520  
Db 2255 GTACTGATCAACATAGAGAGCAATCTCGATTTGCAGATCAAGTGTTTCAGGAGCGT 2314  
Qy 2521 ACTGATTGGACATTTCCAAAGCCTGCTCGTAGGTGGCAACCGATAGGTTGTAGAGTGTG 2580  
Db 2315 ACTGATTGGACATTTCCAAAGCCTGCTCGTAGGTGGCAACCGATAGGTTGTAGAGTGTG 2374  
Qy 2581 CAATACACTTCGGTACAAATTTCAACCCCTTGCACACTGCACAGCTTGGTGTGACACGAT 2640  
Db 2375 CAATACACTTCGGTACAAATTTCAACCCCTTGCACACTGCACAGCTTGGTGTGACACGAT 2434  
Qy 2641 CTTCAATTTCTGGCAGCTCTCTGTCTGTCATATCGACAGCCAAACAGAAATCACCTGGGAAT 2700  
Db 2435 CTTCAATTTCTGGCAGCTCTCTGTCTGTCATATCGACAGCCAAACAGAAATCACCTGGGAAT 2494  
Qy 2701 CAATACCAATTTAGCTTGTGACAGAGAGTCTGAGGCAACGAAATCTGGATCAGCGTATT 2760  
Db 2495 CAATACCAATTTAGCTTGTGACAGAGAGTCTGAGGCAACGAAATCTGGATCAGCGTATT 2554  
Qy 2761 TATCAGCAATTAACTAGAATCTTCAAGAGCCAGAGGATCTCAATACTACACAGGCGTG 2820  
Db 2555 TATCAGCAATTAACTAGAATCTTCAAGAGCCAGAGGATCTCAATACTACACAGGCGTG 2614  
Qy 2821 ATGTGTCATTTTGAACCAATCTTTGGCAGAGTAACGAACTGGTTTCTGGACCAAAATA 2880  
Db 2615 ATGTGTCATTTTGAACCAATCTTTGGCAGAGTAACGAACTGGTTTCTGGACCAAAATA 2674  
Qy 2881 TTTTGTACACTTAGGACAGTCTTCTGTCGTAGCCATAGCAGCTACTCCCTGGGCGC 2940  
Db 2675 TTTTGTACACTTAGGACAGTCTTCTGTCGTAGCCATAGCAGCTACTCCCTGGGCGC 2734  
Qy 2941 CTCTGCTTAGCAGCATACACTTAGCACCACCTTTGTTGGGCAACGTAGATGACTTTCTGGG 3000  
Db 2735 CTCTGCTTAGCAGCATACACTTAGCACCACCTTTGTTGGGCAACGTAGATGACTTTCTGGG 2794  
Qy 3001 TAAGGGTACATCTTTTAGTGGAGATGCAAAACAAATTTCTTTGCAACCAAGCACTT 3060  
Db 2795 TAAGGGTACATCTTTTAGTGGAGATGCAAAACAAATTTCTTTGCAACCAAGCACTT 2854  
Qy 3061 TGGCAGGAACACCCAGCATCAGGGAAGTGGAGGAGAAATTCGGGTTCCACAGGAATAT 3120  
Db 2855 TGGCAGGAACACCCAGCATCAGGGAAGTGGAGGAGAAATTCGGGTTCCACAGGAATAT 2914  
Qy 3121 AGAGCCCAACTTTCTCAATAGTCTTTGCAAAACGAGAGCAGACTACACAGGGCAAGTCT 3180  
Db 2915 AGAGCCCAACTTTCTCAATAGTCTTTGCAAAACGAGAGCAGACTACACAGGGCAAGTCT 2974  
Qy 3181 CAACTTGCACCTCTCCGTTAGTCTGAGCTTCATGGAATTTCTGACGTTATCTATAGAGA 3240  
Db 2975 CAACTTGCACCTCTCCGTTAGTCTGAGCTTCATGGAATTTCTGACGTTATCTATAGAGA 3034  
Qy 3241 GATCAATGGCTCTCTTAAGCTTATCTGGCAATTTGCATAGTTCCTCTGGGAAAGGAGCTT 3300  
Db 3035 GATCAATGGCTCTCTTAAGCTTATCTGGCAATTTGCATAGTTCCTCTGGGAAAGGAGCTT 3094  
Qy 3301 CTAACAGAGTGTCTTCAAGAGCATCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTT 3360  
Db 3095 CTAACAGAGTGTCTTCAAGAGCATCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTT 3154  
Qy 3361 TGTCAACATTTTGCAGCAATTTGTGCAATTTGGTTTGAATTTCCATAATCTGTTCCG 3420  
Db 3155 TGTCAACATTTTGCAGCAATTTGTGCAATTTGGTTTGAATTTCCATAATCTGTTCCG 3214  
Qy 3421 TTTTCTGGATAGAGCAAGAGGCAATTTCAATTTCTTGTGAGGAGGCTTAGAAAGCT 3480  
Db 3215 TTTTCTGGATAGAGCAAGAGGCAATTTCAATTTCTTGTGAGGAGGCTTAGAAAGCT 3274  
Qy 3481 CAATTTTGCACATTTCAATACGACTTTCAGAAAGGAGCTTCTTTAGTTTGGATTTCTTCT 3540  
Db 3275 CAATTTTGCACATTTCAATACGACTTTCAGAAAGGAGCTTCTTTAGTTTGGATTTCTTCT 3334

Qy 3541 TAGGTTGTTCTTGGTGTATCTCGCTTTGGCATCTCTTTCTCTTAGTGACCTTTAGGG 3600  
Db 3335 TAGGTTGTTCTTGGTGTATCTCGCTTTGGCATCTCTTTCTCTTAGTGACCTTTAGGG 3394  
Qy 3601 ACTTCATATCCAGGTTTCTTCCACTCGTCAAGCTGCACACGCTACTTGGCACAATCTAA 3660  
Db 3395 ACTTCATATCCAGGTTTCTTCCACTCGTCAAGCTGCACACGCTACTTGGCACAATCTAA 3454  
Qy 3661 CTAATGCAAAATAAATAAGTTCAGACATTCACAGCTATATCTTCTTTGGATTTAGCTT 3720  
Db 3455 CTAATGCAAAATAAATAAGTTCAGACATTCACAGCTATATCTTCTTTGGATTTAGCTT 3514  
Qy 3721 CTGCAAGTTCATCAGCTTCTCTCTTAATTTTAGCGTTCAACAAAATCTTCTGTCATAATA 3780  
Db 3515 CTGCAAGTTCATCAGCTTCTCTCTTAATTTTAGCGTTCAACAAAATCTTCTGTCATAATA 3574  
Qy 3781 ACCGTTTGTATAAGAACCTTCTGGAGCATTCCTTACGATCCCAAGSTGGCTTCCA 3840  
Db 3575 ACCGTTTGTATAAGAACCTTCTGGAGCATTCCTTACGATCCCAAGSTGGCTTCCA 3634  
Qy 3841 TGGCTCTAAGACCTTTTGAATTTGGCCAAAACAGAGTGGCTTCCAAGTGCAGAAAACCA 3900  
Db 3635 TGGCTCTAAGACCTTTTGAATTTGGCCAAAACAGAGTGGCTTCCAAGTGCAGAAAACCA 3694  
Qy 3901 CACTGTTTGTCAACCCACAAATTTCAAGCAGTCTCCATCAATCCAAATTCGATACCCA 3960  
Db 3695 CACTGTTTGTCAACCCACAAATTTCAAGCAGTCTCCATCAATCCAAATTCGATACCCA 3754  
Qy 3961 GCAACTTTTGTAGTGTCTCAGATGTAGCACCTTTATACCAAAACCGTGCAGACGAGAT 4020  
Db 3755 GCAACTTTTGTAGTGTCTCAGATGTAGCACCTTTATACCAAAACCGTGCAGACGAGAT 3814  
Qy 4021 GGTAGACTCCAGTTTGTCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACCA 4080  
Db 3815 GGTAGACTCCAGTTTGTCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACCA 3874  
Qy 4081 GGCCTCAACAGTAATTTAGAGAGTGCAGCCACAAAGTAGTGAATAGACCATTCGGGCGGT 4140  
Db 3875 GGCCTCAACAGTAATTTAGAGAGTGCAGCCACAAAGTAGTGAATAGACCATTCGGGCGGT 3934  
Qy 4141 CAGTAGTCAAGACGCGCAACAAATTTCTAGCAGGGAATCTTTTGACATCTTTCAGAAA 4200  
Db 3935 CAGTAGTCAAGACGCGCAACAAATTTCTAGCAGGGAATCTTTTGACATCTTTCAGAAA 3994  
Qy 4201 GTTCGTTTTCAGTAGTCAATTTGCGGAGCATCAATATGGGGATTATACAGAAAGCAACAG 4260  
Db 3995 GTTCGTTTTCAGTAGTCAATTTGCGGAGCATCAATATGGGGATTATACAGAAAGCAACAG 4054  
Qy 4261 TGGAAAGTCAATCTACCAACTTTTGGGCTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4320  
Db 4055 TGGAAAGTCAATCTACCAACTTTTGGGCTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4114  
Qy 4321 CATTAGTGAACCTTTTCAATCCCGCAGTGGAGAGAAAAGCAGCGATACTAGCAT 4380  
Db 4115 CATTAGTGAACCTTTTCAATCCCGCAGTGGAGAGAAAAGCAGCGATACTAGCAT 4174  
Qy 4381 TAGCGGCAAGGATGCAACTTTTATCAACAGGGTCTCTATAGATAACCTTAGCGCTGGGA 4440  
Db 4175 TAGCGGCAAGGATGCAACTTTTATCAACAGGGTCTCTATAGATAACCTTAGCGCTGGGA 4234  
Qy 4441 TCATCTTTTGGACAACTTTTCTTTCGCAAAATCTAGGTCGCAAAATCTTCTGATACCAT 4500  
Db 4235 TCATCTTTTGGACAACTTTTCTTTCGCAAAATCTAGGTCGCAAAATCTTCTGATACCAT 4294  
Qy 4501 TATTGTACAATCTTGCAGCACTTCTGATCAGCTCCTCAAAATTTGGTCTCTGTAAACGGATG 4560  
Db 4295 TATTGTACAATCTTGCAGCACTTCTGATCAGCTCCTCAAAATTTGGTCTCTGTAAACGGATG 4354  
Qy 4561 ACTCAATCTGCATTAATTTAGAGTCAAGCTCAGTTCGATTCAGTGAATTCAGGTTGTGCA 4620  
Db 4355 ACTCAATCTGCATTAATTTAGAGTCAAGCTCAGTTCGATTCAGTGAATTCAGGTTGTGCA 4414

4621 QY GCTGTCAGCAGCATAGGGAACAACGCGTTTTCCTACCAAACTCAAGAAATTTATCAAACT 4680  
4415 Db GCTGTCAGCAGCATAGGGAACAACGCGTTTTCCTACCAAACTCAAGAAATTTATCAAACT 4474  
4681 QY CTGCAACACTTTCGCTATCGAGTACGAGGAATGTCATCTGAGTGGACAGTGAG 4740  
4475 Db CTGCAACACTTTCGCTATCGAGTACGAGGAATGTCATCTGAGTGGACAGTGAG 4534  
4741 QY TGTAGTCTTGAGAAATCTGAAAGCCGTATTTTATTATCACTGAGTCACTCAGGAGA 4800  
4535 Db TGTAGTCTTGAGAAATCTGAAAGCCGTATTTTATTATCACTGAGTCACTCAGGAGA 4594  
4801 QY TCCTCTACGCGGAGCGCATCTGTGGCGA----- 4828  
4595 Db TCCTCTACGCGGAGCGCATCTGTGGCGA----- 4828  
4829 QY----- 4828  
4655 Db CTCGTGAAGAGGTGTTGCTGACTCATACAGGCGCTGAATCGCCCCATCATCAGCCAGA 4714  
4829 QY----- 4828  
4715 Db AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGTGGACCACTGTTGTTGATTTGA 4774  
4829 QY----- 4828  
4775 Db ACTTTTGTCTTCCACGGAACGGTCTGGTGTGTCGGAAGATGCGTGATCTGATCCTCA 4834  
4829 QY----- 4828  
4835 Db ACTCAGCAAAAGTTCGATTTATTCAAACAAAGCCGCTCCGCTCAAGTCAAGCGTAATGCT 4894  
4829 QY----- 4828  
4895 Db CTGCCAGTGTTAACAAACCAATTAACCAATCTGATTAGAAAACTCATCGAGCATCAATG 4954  
4829 QY----- 4828  
4955 Db AAACCTGCAATTTATCATATCAGGATATCAATACCATATTTTGTGAAAAGCCGTTCTG 5014  
4829 QY----- 4828  
5015 Db TAATGAGGAGAAAACTCACCGAGGCGAGTTCATAGGATGGCAAGATCCTGGTATCGGTC 5074  
4829 QY----- 4828  
5075 Db TGGGATCCGACTCGTCCAAATCAATACACCTATTAATTCCTCCCTCGTCAAAAATAAG 5134  
4829 QY----- 4828  
5135 Db GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGGCAAAAGCTT 5194  
4829 QY----- 4828  
5195 Db ATGCATTTCTTCCAGACTTGTTCACAGCCGAGCCCATTAAGTTCCTCGTCAAAAATAAG 5254  
4829 QY----- 4828  
5255 Db CGCATCAACCAACCGTTTATTCATTGCTGATTGCGCTGAGCGAGAGCAAAATACGCGATC 5314  
4829 QY----- 4828  
5315 Db GCTGTTAAAGGACAAATTAACAAAGGAATCGAAATGCAACCGGCGAGGAACACTGCGCAG 5374  
4829 QY----- 4828  
5375 Db CGCATCAACAAATTTTCAACCTGAATCAGGATATTCCTTAATACCTGGAATGCTGTTTT 5434  
4829 QY----- 4828  
5435 Db CCCGGGATCGAGTGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGAT 5494  
4829 QY----- 4828

5495 Db GCTCGGAAGAGGCATAAAATTCGGTCAGCCAGTTTGTAGTGTGACCATCTCATCTGTAAACATC 5554  
4829 QY----- 4828  
5555 Db ATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGCGCATCGGCTTCCCATATA 5614  
4829 QY----- 4828  
5615 Db CAATCGATAGATTGTGCAACCTGATTGCCGACATTTATCGGAGCCCATTTATACCCATA 5674  
4829 QY----- 4828  
5675 Db TAAATCAGCATCCATGTTGGAAATTTAATCGGGGCTCGAGCAAGACGTTTCCCGTTGNAAT 5734  
4829 QY----- 4828  
5735 Db ATGGCTCATACACACCCCTTGTTATTACTGTTTATGTAAGACACAGTTTATTGTTTCATGA 5794  
4829 QY----- 4828  
5795 Db TGATATATTTTATCTTGTGCAATGTAAATCATCAGAGATTTTGAGACACAACGTCGCTTTC 5854  
4829 QY----- 4828  
5855 Db CCCCCCCCCCTGCAAGTTCGTCATCACCGCGCCACAGTTCGGTTCGTCGGGCTATAT 5914  
4880 QY CCGCAGCATACCGATGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGGCTTGTTT 4939  
5915 Db CCGCAGCATACCGATGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGGCTTGTTT 5974  
4940 QY CCGGCTGGTATGTTGGCAGGCCCCGCTGGCGGGGACTGTTGGCGGCCATCTCTTTCGA 4999  
5975 Db CCGGCTGGTATGTTGGCAGGCCCCGCTGGCGGGGACTGTTGGCGGCCATCTCTTTCGA 6034  
5000 QY TGCACATTCCTTGGCGGGGCTGCTCAACGGGCTCAACCTTACTTCTGGGCTGCTTCT 5059  
6035 Db TGCACATTCCTTGGCGGGGCTGCTCAACGGGCTCAACCTTACTTCTGGGCTGCTTCT 6094  
5060 QY AATGAGAGTGCATTAAGGAGAGCGTCGAGTATCTATGATTTGAGATATGGAATGCT 5119  
6095 Db AATGAGAGTGCATTAAGGAGAGCGTCGAGTATCTATGATTTGAGATATGGAATGCT 6154  
5120 QY GATACCCGCATTTCTTCACTGTCTTGTAGGTCTCTATCATAGATTATGCCAACTAAAGCAAC 5179  
6155 Db GATACCCGCATTTCTTCACTGTCTTGTAGGTCTCTATCATAGATTATGCCAACTAAAGCAAC 6214  
5180 QY CGGAGAGAGATTTTCATGTAATTTCTCTGACTTTTGTGTCATCAGTAGACTCGAATCG 5239  
6215 Db CGGAGAGAGATTTTCATGTAATTTCTCTGACTTTTGTGTCATCAGTAGACTCGAATCG 6274  
5240 QY TGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTTGGAGACAGTAAATGAATGCC 5299  
6275 Db TGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTTGGAGACAGTAAATGAATGCC 6334  
5300 QY ACCAATAAGAAATCCTTGTATCATAGGAACAACTTCTTGTTCGAACTTTTCGGTGCC 5359  
6335 Db ACCAATAAGAAATCCTTGTATCATAGGAACAACTTCTTGTTCGAACTTTTCGGTGCC 6394  
5360 QY TTGAACTATAAATGTAGAGTGGATATCTCGGTAGGAATGGAGCGGCAAAATGCTTACC 5419  
6395 Db TTGAACTATAAATGTAGAGTGGATATCTCGGTAGGAATGGAGCGGCAAAATGCTTACC 6454  
5420 QY TTCTCGAATCTCAAGAGTATGTAGGTTTGTAGTACTGATGCCAACTTCAGTGACAAAC 5479  
6455 Db TTCTCGAATCTCAAGAGTATGTAGGTTTGTAGTACTGATGCCAACTTCAGTGACAAAC 6514  
5480 QY GTTGTCTATTTCTTCAAAACCAATTCGAAATCCAGAGAAATCAAAGTTGTTTCTTACTATT 5539  
6515 Db GTTGTCTATTTCTTCAAAACCAATTCGAAATCCAGAGAAATCAAAGTTGTTTCTTACTATT 6574  
5540 QY GATCAAGCAGTGGCTTTGAAACTGCAATATGTTGCTGTTTGGAGTCACTTT 5599

Db 6575 GATCAAGCCAGTGGGTCTTGAAGTCAATAGTGTCTGTTGTTTGAAGTCAATCTT 6634  
Qy 5600 TGTATGAATAAATCTAGTCTTTTGAATCTAAATCTTTGACGAGCAAGGCGATATAATACC 5659  
Db 6635 TGTATGAATAAATCTAGTCTTTTGAATCTAAATCTTTGACGAGCAAGGCGATATAATACC 6694  
Qy 5660 CAATCTAAATCTTTTAAACGTTAAAGCAAGATATGTCTGCTGTATTAACCC 5719  
Db 6695 CAATCTAAATCTTTTAAACGTTAAAGCAAGATATGTCTGCTGTATTAACCC 6754  
Qy 5720 AAATCAGCTCGTGTCTGATCCTCATCAATCTTGAAGGCGCTATCTTTTGAAGAAAT 5779  
Db 6814 AAATCAGCTCGTGTCTGATCCTCATCAATCTTGAAGGCGCTATCTTTTGAAGAAAT 6874  
Qy 5780 TTGGGAGATCGATATCGAGAAAAGTACGCTGATTTTAAAGCTGAATTTATCTCAA 5839  
Db 6815 TTGGGAGATCGATATCGAGAAAAGTACGCTGATTTTAAAGCTGAATTTATCTCAA 6874  
Qy 5840 GATCTCTCCCTCGCGGTTCGGTGATGACCGTGAAACCTCTGACACATCGACCTCCG 5899  
Db 6934 GATCTCTCCCTCGCGGTTCGGTGATGACCGTGAAACCTCTGACACATCGACCTCCG 6994  
Qy 5900 GAGACGGTCAACAGTGTCTGTAAGCGGATCGCGGAGCAGACAAAGCCGTCAGGCGCG 5959  
Db 6935 GAGACGGTCAACAGTGTCTGTAAGCGGATCGCGGAGCAGACAAAGCCGTCAGGCGCG 6994  
Qy 5960 TCAGCGGTGTGGCGGTGTGGCGGCGCAGCCATGACCCAGTCAAGTACGATAGCGGA 6019  
Db 6995 TCAGCGGTGTGGCGGTGTGGCGGCGCAGCCATGACCCAGTCAAGTACGATAGCGGA 7054  
Qy 6020 GTGTATCTGGCTTAACCTATCGGCATCAGACGAGATGTACTGAGAGTGACCATATGC 6079  
Db 7055 GTGTATCTGGCTTAACCTATCGGCATCAGACGAGATGTACTGAGAGTGACCATATGC 7114  
Qy 6080 GGTGTAAATACCGCACAGATCGCTAAGGAGAAATACCGCATCAGCGCTCTTCCGCTT 6139  
Db 7115 GGTGTAAATACCGCACAGATCGCTAAGGAGAAATACCGCATCAGCGCTCTTCCGCTT 7174  
Qy 6140 CCTCGCTCACTGACTCGTGTGGCTCGGTTCGGCTGCGGAGCGGTATCAGCTCACT 6199  
Db 7175 CCTCGCTCACTGACTCGTGTGGCTCGGTTCGGCTGCGGAGCGGTATCAGCTCACT 7234  
Qy 6200 CAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACCGAGAAAGAACATGTGAG 6259  
Db 7235 CAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACCGAGAAAGAACATGTGAG 7294  
Qy 6260 CAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACCGAGAAAGAACATGTGAG 6319  
Db 7295 CAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACCGAGAAAGAACATGTGAG 7354  
Qy 6320 GGCTCGGCCCTCGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 6379  
Db 7355 GGCTCGGCCCTCGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 7414  
Qy 6380 CGACGAGCTATAAGATACGAGGCTTTCCGCTGGAAGTTCCTCGTGGCTCTCTG 6439  
Db 7415 CGACGAGCTATAAGATACGAGGCTTTCCGCTGGAAGTTCCTCGTGGCTCTCTG 7474  
Qy 6440 TTTCCGACCTCGCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGAGCGGTGGCG 6499  
Db 7475 TTTCCGACCTCGCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGAGCGGTGGCG 7534  
Qy 6500 TTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGCTCCAGCTGG 6559  
Db 7535 TTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGCTCCAGCTGG 7594  
Qy 6560 GCTGTGTGACGAAACCCCGGTTACGCGGACCGCTGCGCTTATCCCGTAACTATCGTC 6619  
Db 7595 GCTGTGTGACGAAACCCCGGTTACGCGGACCGCTGCGCTTATCCCGTAACTATCGTC 7654  
Qy 6620 TTTGAGTCCAAACCCCGGTTAAGACACGCTTATCGCCACTGGCAGCGCTATGTAACAGGA 6679  
Db 7655 TTTGAGTCCAAACCCCGGTTAAGACACGCTTATCGCCACTGGCAGCGCTATGTAACAGGA 7714

Qy 6680 TTAGCAGACGAGGATGTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCCTAATACG 6739  
Db 7715 TTAGCAGACGAGGATGTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCCTAATACG 7774  
Qy 6740 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTCGGA 6799  
Db 7775 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTCGGA 7834  
Qy 6800 AAAGAGTTCGTAGCTCTTGATCCGCAACAAACACCGCTGCTAGCGGTGTTTTTTG 6859  
Db 7835 AAAGAGTTCGTAGCTCTTGATCCGCAACAAACACCGCTGCTAGCGGTGTTTTTTG 7894  
Qy 6860 TTTGCAAGCAGAGATAGCGCGAAGAAAGGATCTCAAGAGATCCTTTGATCTTTT 6919  
Db 7895 TTTGCAAGCAGAGATAGCGCGAAGAAAGGATCTCAAGAGATCCTTTGATCTTTT 7954  
Qy 6920 CTACGGGTCTGACGCTCAGTGAACGAAACCTCAAGTAAAGGATTTTGGTCAAGAT 6979  
Db 7955 CTACGGGTCTGACGCTCAGTGAACGAAACCTCAAGTAAAGGATTTTGGTCAAGAT 8014  
Qy 6980 TATCAAAAAGGATCTTACCTAGATCTCTTTTAAATGAAAGTTTTAAATCAATCT 7039  
Db 8015 TATCAAAAAGGATCTTACCTAGATCTCTTTTAAATGAAAGTTTTAAATCAATCT 8074  
Qy 7040 AAGATATATAGTAAACCTTGGTCTGACAGTTACCAATGCTTAATCACTGAGGACCTA 7099  
Db 8075 AAGATATATAGTAAACCTTGGTCTGACAGTTACCAATGCTTAATCACTGAGGACCTA 8134  
Qy 7100 TCTCAGCGATCTCTTAATTCGTTTCATCATAGTTGCTGACTCCCGCTGCTAGATAA 7159  
Db 8135 TCTCAGCGATCTCTTAATTCGTTTCATCATAGTTGCTGACTCCCGCTGCTAGATAA 8194  
Qy 7160 CTACGATACGGGAGGCTTACCATCTGCGCCAGTGTCTCAATGATACCGCAGACCCAC 7219  
Db 8195 CTACGATACGGGAGGCTTACCATCTGCGCCAGTGTCTCAATGATACCGCAGACCCAC 8254  
Qy 7220 GCTCAGCGGTCCAGATTTATCAGCAATAAACCAGCAGCCGGAAGGCGAGCGAGAA 7279  
Db 8255 GCTCAGCGGTCCAGATTTATCAGCAATAAACCAGCAGCCGGAAGGCGAGCGAGAA 8314  
Qy 7280 GTGGTCTTGCACTTTATCGGCTCATCCAGTCTAATTAATTTGTCGGGAGCTAGAG 7339  
Db 8315 GTGGTCTTGCACTTTATCGGCTCATCCAGTCTAATTAATTTGTCGGGAGCTAGAG 8374  
Qy 7340 TAAAGTATTCGCCAGTAAATAGTTTTCGCAACGTTGTCGCAATGCTGAGGATCGTG 7399  
Db 8375 TAAAGTATTCGCCAGTAAATAGTTTTCGCAACGTTGTCGCAATGCTGAGGATCGTG 8434  
Qy 7400 TGTCACTGCTGCTGTTGATGGCTTCATTCAGCTCCGCTTCCCAACGATCAAGCGAG 7459  
Db 8435 TGTCACTGCTGCTGTTGATGGCTTCATTCAGCTCCGCTTCCCAACGATCAAGCGAG 8494  
Qy 7460 TTAATGATCCCGCATGTTGCAAAAAGCGGTAGCTCTTTCGCTCCTCGATCGTTG 7519  
Db 8495 TTAATGATCCCGCATGTTGCAAAAAGCGGTAGCTCTTTCGCTCCTCGATCGTTG 8554  
Qy 7520 TCAGAAATAGTTGGCGCAGTGTATCACTCATGTTATGCGAGCACTGCATTAATCTC 7579  
Db 8555 TCAGAAATAGTTGGCGCAGTGTATCACTCATGTTATGCGAGCACTGCATTAATCTC 8614  
Qy 7580 TTAATGATCCCGCATGTTGCAAAAAGCGGTAGCTCTTTCGCTCCTCGATCGTTG 7639  
Db 8615 TTAATGATCCCGCATGTTGCAAAAAGCGGTAGCTCTTTCGCTCCTCGATCGTTG 8674  
Qy 7640 TCTGAGAAATAGTGTGCGCGCAGTGTCTTTCGCGCGGTCAACACGGGATAATA 7699  
Db 8675 TCTGAGAAATAGTGTGCGCGCAGTGTCTTTCGCGCGGTCAACACGGGATAATA 8734  
Qy 7700 CCGCGCACATAGCAGAACTTTAAAGTGTCTCATCTTTGAAAGCGTTCTTCGGGCGGAA 7759  
Db 8735 CCGCGCACATAGCAGAACTTTAAAGTGTCTCATCTTTGAAAGCGTTCTTCGGGCGGAA 8794

```

QY 7760 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 7819
DB 8795 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 8854
QY 7820 ACTGATCTTCAGATCTTTTACTTTCACGAGCTTTCTGGGTGAGCAAAACAGGAGGC 7879
DB 8855 ACTGATCTTCAGATCTTTTACTTTCACGAGCTTTCTGGGTGAGCAAAACAGGAGGC 8914
QY 7880 AABATGCGCAAAAGGAAATAGGCGCACACGGAATGTTGAATCTCATCTCTTCC 7939
DB 8915 AAAATGCGCAAAAGGAAATAGGCGCACACGGAATGTTGAATCTCATCTCTTCC 8974
QY 7940 TTTTTCATATATTGAAGCATTTATCAGGGTTATTCTCATGAGCGGATACATATTG 7999
DB 8975 TTTTTCATATATTGAAGCATTTATCAGGGTTATTCTCATGAGCGGATACATATTG 9034
QY 8000 AATGATATTAGAAAATARAACAAATAGGGTTCGCGCACATTTCCCGAAGTCCAC 8059
DB 9035 AATGATATTAGAAAATARAACAAATAGGGTTCGCGCACATTTCCCGAAGTCCAC 9094
QY 8060 CTGACGCTTAAGAAACCATTTATCATGACATTAACCTATAAAATAGCGGTATCACGA 8119
DB 9095 CTGACGCTTAAGAAACCATTTATCATGACATTAACCTATAAAATAGCGGTATCACGA 9154
QY 8120 GGCCCTTCGCTTCAAGAAATTAATCTCATGTTGACAGTTATCATCATAGCTGAC 8179
DB 9155 GGCCCTTCGCTTCAAGAAATTAATCTCATGTTGACAGTTATCATCATAGCTGAC 9214
QY 8180 TCATGTTGTTATTTGAATAGACGCGAGATCGGGAACACTGAAATAAACAAGTTATT 8239
DB 9215 TCATGTTGTTATTTGAATAGACGCGAGATCGGGAACACTGAAATAAACAAGTTATT 9274
QY 8240 CG.8241
DB 9275 CG.9276

RESULT 10
AAL60801
ID AAL60801 standard; DNA; 9276 BP.
AC AAL60801;
XX
DT 03-SEP-2003 (first entry)
XX
DE Pichia pastoris vector pPIC9K DNA.
XX
KW Serine protease 17; CVSP17; tumour; cancer; antisense therapy; prostate;
XX breast; cervix; lung; ovary; colon; gene therapy; pPIC9K vector; ds.
XX
OS Pichia pastoris.
XX
FH Key Location/Qualifiers
FT promoter 1..948
FT primer_bind /tag= a
FT /note= "5' AOX1 promoter fragment"
FT 855..875
FT /tag= v
FT /bound moiety= "AOX1 primer"
FT 949..1218
FT /tag= b
FT /note= "Alpha-factor secretion signal"
FT 1152..1172
FT /tag= v
FT /bound moiety= "Alpha-factor primer"
FT 1192..1241
FT /tag= b
FT /note= "Multiple cloning site"
FT 1253..1586
FT /tag= v
FT /note= "AOX1 transcription termination region"
FT 1327..1347
FT /tag= v

```

```

FT primer_bind /bound moiety= "AOX1 primer"
FT complement(1980..4514)
FT /tag= v
FT /note= "HIS4 ORF"
FT misc_feature complement(4928..5743)
FT /tag= v
FT /note= "Kanamycin resistance gene"
FT promoter 6122..6879
FT /tag= a
FT /note= "3' AOX1 promoter fragment"
FT misc_feature complement(7288..7961)
FT /tag= v
FT /note= "ColE1 origin"
FT misc_feature complement(8106..8966)
FT /tag= v
FT /note= "Ampicillin resistance gene"
XX WO2003044179-A2.
XX
XX 30-MAY-2003.
XX
XX 20-NOV-2002; 2002WO-US037626.
XX
XX 20-NOV-2001; 2001US-0332015P.
XX (CORV-) CORVAS INT INC.
XX
XX Madison EL, Ong EO;
XX
XX WPI; 2003-449816/42.
XX
XX New substantially purified serine protease 17 polypeptide and encoding
XX nucleic acid, useful for diagnosing and treating tumor conditions and/or
XX cancer, particularly of the breast, cervix, prostate, lung, ovary or
XX colon.
XX
XX Example 2; Page 185-188; 189pp; English.
XX
XX The invention relates to serine protease 17 polypeptide designated CVSP17
XX and its corresponding nucleic acid sequence. The invention also relates
XX to a method using CVSP17 protein to identify compounds that modulate its
XX protease activity. The method is useful for preventing, diagnosing and
XX treating disorders related to the serine protease 17 activity, such as
XX tumour conditions and/or cancer, particularly of the breast, prostate,
XX cervix, lung, ovary or colon. CVSP17 DNA is used in gene therapy and in
XX antisense therapy. The present sequence is Pichia pastoris vector pPIC9K
XX DNA. This sequence is used to illustrate the method of the invention
XX
XX SQ Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;
XX
XX Query Match 79.6%; Score 6561.6; DB 7; Length 9276;
XX Best Local Similarity 84.7%; Pred. NO. 0;
XX Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;
XX
XX QY 1 AGATCTAATCAATCCAAAGACGAAAGTTGAATGAACACCTTTTGGCATCCGACATCCACAG 60
XX |||||
XX Db 1 AGATCTAATCAATCCAAAGACGAAAGTTGAATGAACACCTTTTGGCATCCGACATCCACAG 60
XX |||||
XX QY 61 GTCCATTCTCACATATAAGTCCAAACGCAACGAGGAGGGATACACTAGCAGCAGACCGT 120
XX |||||
XX Db 61 GTCCATTCTCACATATAAGTCCAAACGCAACGAGGAGGGATACACTAGCAGCAGACCGT 120
XX |||||
XX QY 121 TCGAAACGAGGACCTCCACCTCTTCTCTCAACACCCACTTTTGGCATCCGAAAAACC 180
XX |||||
XX Db 121 TCGAAACGAGGACCTCCACCTCTTCTCTCAACACCCACTTTTGGCATCCGAAAAACC 180
XX |||||
XX QY 181 AGCCAGATTATGGGCTTGAATGGAGCTCGCTCATTCCTTCTTATTAGCTACTA 240
XX |||||
XX Db 181 AGCCAGATTATGGGCTTGAATGGAGCTCGCTCATTCCTTCTTATTAGCTACTA 240
XX |||||
XX QY 241 ACACCATGACTTTATAGCTGTCTATCCTGGCCCCCTGGCGAGGTTCACTGTTGTTA 300
XX |||||
XX Db 241 ACACCATGACTTTATAGCTGTCTATCCTGGCCCCCTGGCGAGGTTCACTGTTGTTA 300
XX |||||

```

QY 301 TTTCGAATGCAACAAAGCTCCGCAATTACACCCGAACATCACTCCAGATGAGGGCTTTCTG 360  
DB 301 TTTCGAATGCAACAAAGCTCCGCAATTACACCCGAACATCACTCCAGATGAGGGCTTTCTG 360  
QY 361 AGTGTGGGTCAAATAGTTTTCATGTTCCCAAAATGGCCCAAACTGACAGCTTTAAACGCT 420  
DB 361 AGTGTGGGTCAAATAGTTTTCATGTTCCCAAAATGGCCCAAACTGACAGCTTTAAACGCT 420  
QY 421 GTCTTGGAACTAATATGACAAAAGCGTGATCTCATCCAGATGAACCTAAGTTGGTTG 480  
DB 421 GTCTTGGAACTAATATGACAAAAGCGTGATCTCATCCAGATGAACCTAAGTTGGTTG 480  
QY 481 TTGAAATGCTTAACCGCGAGTTGTCACAAAAGAAACTTCCAAAAGTCGCCATACCGTTTGT 540  
DB 481 TTGAAATGCTTAACCGCGAGTTGTCACAAAAGAAACTTCCAAAAGTCGCCATACCGTTTGT 540  
QY 541 CTGTGTTGGTATTGATGACGAATGCTCAAAAATAATCTCAATTAATGCTTAGCGCAGTCT 600  
DB 541 CTGTGTTGGTATTGATGACGAATGCTCAAAAATAATCTCAATTAATGCTTAGCGCAGTCT 600  
QY 601 CTCTATCGCTTCTGAAACCCCGGTGCACTGTGCGGAACGCAATGGGAAACACCCGCT 660  
DB 601 CTCTATCGCTTCTGAAACCCCGGTGCACTGTGCGGAACGCAATGGGAAACACCCGCT 660  
QY 661 TTTTGGATGATTATGCAATGTCTCCACATTTGATGCTTCCAAAGATTCTGGTGGGAATACT 720  
DB 661 TTTTGGATGATTATGCAATGTCTCCACATTTGATGCTTCCAAAGATTCTGGTGGGAATACT 720  
QY 721 GCTGATAGCCTAAGTTCATGATCAAAATTAATCTGTTCTAACCCCTACTGTGACGCAAT 780  
DB 721 GCTGATAGCCTAAGTTCATGATCAAAATTAATCTGTTCTAACCCCTACTGTGACGCAAT 780  
QY 781 ATATAACAGAGGAAGCTGCCGTCTTAACCTTTTATTAACCTTTTATCATATTAGCTT 840  
DB 781 ATATAACAGAGGAAGCTGCCGTCTTAACCTTTTATTAACCTTTTATCATATTAGCTT 840  
QY 841 ACTTTTCATAATTCGACTGGTTTCCAAATGCAAGCTTTTGATTTTAAACGACTTTTAAACGA 900  
DB 841 ACTTTTCATAATTCGACTGGTTTCCAAATGCAAGCTTTTGATTTTAAACGACTTTTAAACGA 900  
QY 901 CAACTTGAGAGATCAAAAACAACTAATTAATGGAAGATCCAAACGATGAGATTTCCT 960  
DB 901 CAACTTGAGAGATCAAAAACAACTAATTAATGGAAGATCCAAACGATGAGATTTCCT 960  
QY 961 TCAATTTTACTGCAATTTTATTCGAGATCCTCCGCAATTAGCTGCCAGTCAACACT 1020  
DB 961 TCAATTTTACTGCAATTTTATTCGAGATCCTCCGCAATTAGCTGCCAGTCAACACT 1020  
QY 1021 ACAAAGAGATGAACCGCAAAATTCGGCTGAAGCTGTCACTCGGTTACTCAGATTTA 1080  
DB 1021 ACAAAGAGATGAACCGCAAAATTCGGCTGAAGCTGTCACTCGGTTACTCAGATTTA 1080  
QY 1081 GAAGGGGATTCGATGTTGCTGTTTGCCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140  
DB 1081 GAAGGGGATTCGATGTTGCTGTTTGCCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140  
QY 1141 TTTATAAATFACTATTGCCAGCATGCTGCTAAAGAAGAGGGGTATCTCTGAGAAA 1200  
DB 1141 TTTATAAATFACTATTGCCAGCATGCTGCTAAAGAAGAGGGGTATCTCTGAGAAA 1200  
QY 1201 AGAGAGGCTGAAGCCCGACCGCCAGCGCCAGGCCCCAGGAGGAACAATGATCATG 1260  
DB 1201 AGAGAGGCTGAAGCTTA  
QY 1261 GCCCCCCGGAGAGGATTAATTTGGCTTCCCGGTGTCACCGCCAGCAGTGCACGGAG 1320  
DB 1218 ----- 1217  
QY 1321 AGAGGTTGCTGTTTGTGATGACAGTGTCCGGGATTTCCCGTGGTGTTCACCCCATGGCC 1380  
DB 1218 ----- 1217

QY 1381 ATCGAAGACACTCAAGAGAAGAAATGTCCCTTCTAACTAGTGGCGTAGAATTCCTTAGG 1440  
DB 1218 ----- CGTAGAATTTCCCTTAGG 1234  
QY 1441 CGGCCCGGAATTAATTCGCTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACCTTAG 1500  
DB 1235 CGGCCCGGAATTAATTCGCTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACCTTAG 1294  
QY 1501 AGAAGACCGCTCTTCTGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTCCTCGTAG 1560  
DB 1295 AGAAGACCGCTCTTCTGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTCCTCGTAG 1354  
QY 1561 ATCGAGCTTCAATTTTGTACTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1620  
DB 1355 ATCGAGCTTCAATTTTGTACTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1414  
QY 1621 GTCAATTTGTTTCTTCTGCTAGCAGCTTCTCTGATCAGCCTATCTCCAGCTGATGAA 1680  
DB 1415 GTCAATTTGTTTCTTCTGCTAGCAGCTTCTCTGATCAGCCTATCTCCAGCTGATGAA 1474  
QY 1681 TATCTTGTGTTAGGGTTTGGGAAATCAATTCGAGTTTTCATGTTTCTTGGTATTTCCT 1740  
DB 1475 TATCTTGTGTTAGGGTTTGGGAAATCAATTCGAGTTTTCATGTTTCTTGGTATTTCCT 1534  
QY 1741 ACTCCTCTTTCAGAGTACAGAATTAAGTGAAGTTCTGTTGCAAGCTTATCGATAA 1800  
DB 1535 ACTCCTCTTTCAGAGTACAGAATTAAGTGAAGTTCTGTTGCAAGCTTATCGATAA 1594  
QY 1801 GCTTTAATGGGTAGTTTATCAGTTAAATGCTTAACGAGTTCAGGACCGCTGATGAA 1860  
DB 1595 GCTTTAATGGGTAGTTTATCAGTTAAATGCTTAACGAGTTCAGGACCGCTGATGAA 1654  
QY 1861 ATCTAACAATGGCTCATCTCGGACCGCTCACCTCGGATGCTGTAGGCATAGG 1920  
DB 1655 ATCTAACAATGGCTCATCTCGGACCGCTCACCTCGGATGCTGTAGGCATAGG 1714  
QY 1921 CTTGTTATGCGGTACTGCGGGCTCTTGGGGATTCGTCATTCGACAGCATCCG 1980  
DB 1715 CTTGTTATGCGGTACTGCGGGCTCTTGGGGATTCGTCATTCGACAGCATCCG 1774  
QY 1981 CAGTCACTAGGGCTGCTGCTAGCGCTATATGGTTGATGCAATTTCTATCGCACCCGT 2040  
DB 1775 CAGTCACTAGGGCTGCTGCTAGCGCTATATGGTTGATGCAATTTCTATCGCACCCGT 1834  
QY 2041 TCTCGGAGCATCTGCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTTCTGCTACTGG 2100  
DB 1835 TCTCGGAGCATCTGCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTTCTGCTACTGG 1894  
QY 2101 AGCCACTATCGACTAGCGCATCATGCGGACCAACCCGCTCTGTGATCTATCGAATCTA 2160  
DB 1895 AGCCACTATCGACTAGCGCATCATGCGGACCAACCCGCTCTGTGATCTATCGAATCTA 1954  
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCCAAGTTCTCCATACGAACCTTAA 2220  
DB 1955 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCCAAGTTCTCCATACGAACCTTAA 2014  
QY 2221 CAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCAGCTTGGCCAA 2280  
DB 2015 CAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCAGCTTGGCCAA 2074  
QY 2281 TATGTTTTCAGTCCCTCAGGAGTTTACGCTCTTGTGAAGTGAATGAACTTCTGGAAGTTGCAG 2340  
DB 2075 TATGTTTTCAGTCCCTCAGGAGTTTACGCTCTTGTGAAGTGAATGAACTTCTGGAAGTTGCAG 2134  
QY 2341 TGTAACTCCGCTGATATTGACGGGCATATCCGTTACGTTGGCAAGTGTGGTTCGACCGG 2400  
DB 2135 TGTAACTCCGCTGATATTGACGGGCATATCCGTTACGTTGGCAAGTGTGGTTCGACCGG 2194  
QY 2401 AGGAGTAATCTCCCAACTCTCTGGAGAGTAGGACCAACAAACAGATCCAGCGTGT 2460  
DB 2195 AGGAGTAATCTCCCAACTCTCTGGAGAGTAGGACCAACAAACAGATCCAGCGTGT 2254  
QY 2461 GTACTTTGATCAACATAAGAGAAGCAATTCCTCGATTTTGCAGGATCAAGTGTTCAGGACGT 2520

DB 2255 GTACTTGATCAACATAGAGAGACATCTCGATTTCCAGGATCAAGTGTTCAGGAGCGT 2314  
QY 2521 ACTGATTGGACATTTCCAAAGCCCTGCTCGTAGGTTGCAACCGATAGGGTTGTAGAGTGTG 2580  
DB 2315 ACTGATTGGACATTTCCAAAGCCCTGCTCGTAGGTTGCAACCGATAGGGTTGTAGAGTGTG 2374  
QY 2581 CAATACACTTTCGCTGACAAATTTCAACCCCTTGGCAAGCTGCAAGCTTGTGTAACAGCAT 2640  
DB 2375 CAATACACTTTCGCTGACAAATTTCAACCCCTTGGCAAGCTGCAAGCTTGTGTAACAGCAT 2434  
QY 2641 CTTCAATTTCTGGCAAGCTCCTTGTCTGCTCATATCGACAGCAACAGAAATCACCTGGGAAT 2700  
DB 2435 CTTCAATTTCTGGCAAGCTCCTTGTCTGCTCATATCGACAGCAACAGAAATCACCTGGGAAT 2494  
QY 2701 CAATACCAATGTTGAGCTTGACAGAGAGGCTGAGGCAACGAAATCTGGATCAGCGTATT 2760  
DB 2495 CAATACCAATGTTGAGCTTGACAGAGAGGCTGAGGCAACGAAATCTGGATCAGCGTATT 2554  
QY 2761 TATCAGCAATAAATAGAACTTTCAAGAGGCCCGCAGCAGCATGTCAATACTACACAGGGCTG 2820  
DB 2555 TATCAGCAATAAATAGAACTTTCAAGAGGCCCGCAGCAGCATGTCAATACTACACAGGGCTG 2614  
QY 2821 ATGTGTCATTTGAAACCAATCATCTTGGCAGCAGTAACGAACTGTTTCCTGGACCAATA 2880  
DB 2615 ATGTGTCATTTGAAACCAATCATCTTGGCAGCAGTAACGAACTGTTTCCTGGACCAATA 2674  
QY 2881 TTTTGTACACTTAGGACAGTTCCTGTCGTAAGCCATAGCAGTACTGCTGGGCGC 2940  
DB 2675 TTTTGTACACTTAGGACAGTTCCTGTCGTAAGCCATAGCAGTACTGCTGGGCGC 2734  
QY 2941 CTCCTGCTAGCAGATACACTTAGCACCACCTTGTGGGCAAGTAGACTTCTGGGG 3000  
DB 2735 CTCCTGCTAGCAGATACACTTAGCACCACCTTGTGGGCAAGTAGACTTCTGGGG 2794  
QY 3001 TAAAGGTACCACTCTCTAGTGGAGATGCAAAACAAATTTCTTTCGACACGACCACTT 3060  
DB 2795 TAAAGGTACCACTCTCTAGTGGAGATGCAAAACAAATTTCTTTCGACACGACCACTT 2854  
QY 3061 TGGCAGGAACCCAGCATCAGGAAAGTGAAGCAGAAATGGCGTTCCACAGGAAATAT 3120  
DB 2855 TGGCAGGAACCCAGCATCAGGAAAGTGAAGCAGAAATGGCGTTCCACAGGAAATAT 2914  
QY 3121 AGAGGCCAACTTCTCAATAGTCTTGCAAAACGAGAGCAGATACACACAGGCGAAGTCT 3180  
DB 2915 AGAGGCCAACTTCTCAATAGTCTTGCAAAACGAGAGCAGATACACACAGGCGAAGTCT 2974  
QY 3181 CAACCTTGCAACGCTCTCCGTTAGTTGAGCTTCATGGAATTTCCAGCTTATCTATAGAGA 3240  
DB 2975 CAACCTTGCAACGCTCTCCGTTAGTTGAGCTTCATGGAATTTCCAGCTTATCTATAGAGA 3034  
QY 3241 GATCAATGGCTCTCTTAAGCTTATCTGGCAATGCAATGCTCTCTGGGAAGGAGCTT 3300  
DB 3035 GATCAATGGCTCTCTTAAGCTTATCTGGCAATGCAATGCTCTCTGGGAAGGAGCTT 3094  
QY 3301 CTACACAGGCTCTCTCAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTAAAGAGGCTT 3360  
DB 3095 CTACACAGGCTCTCTCAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAGAGGCTT 3154  
QY 3361 TGTCAACATTTGACGAACATGTCGAATTTGGTTTGTGACTTAATTCATATCTGTTCCG 3420  
DB 3155 TGTCAACATTTGACGAACATGTCGAATTTGGTTTGTGACTTAATTCATATCTGTTCCG 3214  
QY 3421 TTTTCTGATAGGACGAGAGGCACTTCAATTTCTTGTGAGAGGCGCTTAGAAGCT 3480  
DB 3215 TTTTCTGATAGGACGAGAGGCACTTCAATTTCTTGTGAGAGGCGCTTAGAAGCT 3274  
QY 3481 CAATTTGCAAAATTCATAGCACTTTCAGAGGAGCTTCTTTAGGTTTGGATTCTTCTT 3540  
DB 3275 CAATTTGCAAAATTCATAGCACTTTCAGAGGAGCTTCTTTAGGTTTGGATTCTTCTT 3334  
QY 3541 TAGGTTGTTCTTGGTGTATCTTGGCTTGGCATCTCTTCTTCTTCTAGTGACCTTTAGGG 3600

DB 3335 TAGGTTGTTCTTGGTGTATCTTGGCTTGGCATCTCTCTTCTTCTTAGTGACCTTTTAGGG 3394  
QY 3601 ACTTCATATCCAGGTTTCTCTCCACCTCGTCCACGCTCACACCGTACTTGGCAGCATCTAA 3660  
DB 3395 ACTTCATATCCAGGTTTCTCTCCACCTCGTCCACGCTCACACCGTACTTGGCAGCATCTAA 3454  
QY 3661 CTAATGCAAAATAAAATAAGTCAGCAGCATTTCCAGGCTTATATCTTCTTGGATTAGCTT 3720  
DB 3455 CTAATGCAAAATAAAATAAGTCAGCAGCATTTCCAGGCTTATATCTTCTTGGATTAGCTT 3514  
QY 3721 CTGCAAGTTCATCAGCTTCTCTCTTAATTTTAGCGTTCAACAAACTTCGTCGTCAAATA 3780  
DB 3515 CTGCAAGTTCATCAGCTTCTCTCTCTTAATTTTAGCGTTCAACAAACTTCGTCGTCAAATA 3574  
QY 3781 ACCGTTTGGTATAGAACTTCTGGAGCATTTCTAGCATTTCCAGTCCCAAGGTGGCTTCCA 3840  
DB 3575 ACCGTTTGGTATAGAACTTCTGGAGCATTTCTAGCATTTCCAGTCCCAAGGTGGCTTCCA 3634  
QY 3841 TGGCTCTAAGACCTTTTGAATTTGGCCAAACAGGAGTGGCTTCCAGTGCAGAGAACCAA 3900  
DB 3635 TGGCTCTAAGACCTTTTGAATTTGGCCAAACAGGAGTGGCTTCCAGTGCAGAGAACCAA 3694  
QY 3901 CACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCACAATCOAATTCGATACCCA 3960  
DB 3695 CACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCACAATCOAATTCGATACCCA 3754  
QY 3961 GCAACTTTTGTGCTCCAGATGTAGCACCTTTATACCAAAACCGTGCAGCAGGATT 4020  
DB 3755 GCAACTTTTGTGCTCCAGATGTAGCACCTTTATACCAAAACCGTGCAGCAGGATT 3814  
QY 4021 GGTAGACTCCAGTTGCTCTTATAGCTTCGGAATAGACTTTTGTGACGAGTACACCA 4080  
DB 3815 GGTAGACTCCAGTTGCTCTTATAGCTTCGGAATAGACTTTTGTGACGAGTACACCA 3874  
QY 4081 GSCCCCAACGAGTAATTAGAAGAGTCAGCCACCAAGTAGTGAATAGACCATTCGGGGCGGT 4140  
DB 3875 GSCCCCAACGAGTAATTAGAAGAGTCAGCCACCAAGTAGTGAATAGACCATTCGGGGCGGT 3934  
QY 4141 CAGTAGTCAAGACGCCAACAAATTTCACTGACAGGAACTTTTGTGACATCTTCAGAAA 4200  
DB 3935 CAGTAGTCAAGACGCCAACAAATTTCACTGACAGGAACTTTTGTGACATCTTCAGAAA 3994  
QY 4201 GTTCGTATTCAGTAGTCAATTCGCCAGCATCAATTAATGGGATTTATACCAAGCAACAG 4260  
DB 3995 GTTCGTATTCAGTAGTCAATTCGCCAGCATCAATTAATGGGATTTATACCAAGCAACAG 4054  
QY 4261 TGGAGTCAATCTACCAACTTTTGGGCTCTCAGAAAAGCATTAACAGTCTCTACTACCGC 4320  
DB 4055 TGGAGTCAATCTACCAACTTTTGGGCTCTCAGAAAAGCATTAACAGTCTCTACTACCGC 4114  
QY 4321 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAAAGAAAGGACAGCATCTAGCAT 4380  
DB 4115 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAAAGAAAGGACAGCATCTAGCAT 4174  
QY 4381 TAGCGGCAAGGATCAACTTTTATCAACAGGCTCTATAGATAACCCCTAGCGCTTGGGA 4440  
DB 4175 TAGCGGCAAGGATCAACTTTTATCAACAGGCTCTATAGATAACCCCTAGCGCTTGGGA 4234  
QY 4441 TCATCTTTGGACCACTTTCTTCTGCAATCTAGTCCAAATCACTTCAATTCATGATACCAT 4500  
DB 4235 TCATCTTTGGACCACTTTCTTCTGCAATCTAGTCCAAATCACTTCAATTCATGATACCAT 4294  
QY 4501 TATTGTCAAACTTGACGAAGTTGTGATCAGCTCTCTCAAAATTTGGTCTCTCTGTAACGGATG 4560  
DB 4295 TATTGTCAAACTTGACGAAGTTGTGATCAGCTCTCTCAAAATTTGGTCTCTCTGTAACGGATG 4354  
QY 4561 ACTCAACTTGACATTAATTTGAGCTCAGTCGATGAGTGAATGATCAGGTTGTGCA 4620  
DB 4355 ACTCAACTTGACATTAATTTGAGCTCAGTCGATGAGTGAATGATCAGGTTGTGCA 4414  
QY 4621 GCTGCTCAGCAGCATAGGAAACAGCGGCTTTTCTTACCAAACTCAAGGAATTTATCAAACT 4680  
DB 4415 GCTGCTCAGCAGCATAGGAAACAGCGGCTTTTCTTACCAAACTCAAGGAATTTATCAAACT 4474

Qy	4681	CTGCAACACTTGGCTATGTCAGGTAGCAAGGAAATGTCTATCTTGAAGTCGGACAGTGAG	4740	Qy	4829	-----	4828
Dd	4475	CTGCAACACTTGGCTATGTCAGGTAGCAAGGAAATGTCTATCTTGAAGTCGGACAGTGAG	4534	Dd	5555	ATTGGCAACGCTACCTTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGCTTCCCATTA	5614
Qy	4741	TGTAGTCTTTGAGAAATCTTGAAGCCGTAATTTTATTAATCACTAGTCACTCAGTCAAGAGA	4800	Dd	4829	-----	4828
Dd	4535	TGTAGTCTTTGAGAAATCTTGAAGCCGTAATTTTATTAATCACTAGTCACTCAGTCAAGAGA	4594	Qy	5615	CAATCGATAGATTGTGCAACCTGATTGTCGGACATTAATCGCAGGCCCATTTATACCCATA	5674
Qy	4801	TCCTCTACGCGGACGATCTGTGCGCA-----	4828	Qy	4829	-----	4828
Dd	4595	TCCTCTACGCGGACGATCTGTGCGCA-----	4828	Dd	5675	TAAATCAGCATCCATGTTGGAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT	5734
Qy	4829	-----	4828	Qy	4829	-----	4828
Dd	4655	CTCGTGAAGAAGTGTGCTGACTCATACCAGGCTGAATCGCCCATCATCCAGCCAGA	4714	Dd	5735	ATGGCTCATACACCCCTTGTATTAAGTAAAGCAGACAGTTTATTTGTTTCATGA	5794
Qy	4829	-----	4828	Qy	4829	-----	4828
Dd	4715	AAGTAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGTGGAACAGTTGGTGAATTTGA	4774	Dd	5795	TGATATATTTTATCTTGTGCAATGTAAACATCAGAGATTTTGAGACACAAACGTTGCTTC	5854
Qy	4829	-----	4828	Qy	4829	-----	4828
Dd	4775	ACTTTTGCTTTGCCACGGAACGGTCTGGTGTGCGGAAGATGCGTGATCTGATCCTTCA	4834	Dd	5855	CCCCCCCCCTGTCAGGTGCGCATCACCGGCCCAAGGTGCGGTGCTGGCGCCTATAT	5914
Qy	4829	-----	4828	Qy	4880	CGCCGACATCACCGATGGGGAAGATCGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTT	4939
Dd	4835	ACTCAGCAAAAGTTCTGATTTATCAACAAAGCCGCTCCGTCAGGTCAAGTCAAGTATGCT	4894	Dd	5915	CGCCGACATCACCGATGGGGAAGATCGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTT	5974
Qy	4829	-----	4828	Qy	4940	CGCGTGGGTATGTTGGCAGGCCCCCGTGGCGGGGACTGTTGGCGGCCATCTCTCTTGA	6034
Dd	4895	CTGCGAGTTTACAAACCAATTAACCAATTTCTGATTAGAAAACCTCATCGACATCAATG	4954	Qy	5000	TGCACCATTTCTTGGCGCGGCTCAACGGCCCTCAACCTACTACTGGGCTGCTTCT	5059
Qy	4829	-----	4828	Dd	6035	TGCACCATTTCTTGGCGCGGCTCAACGGCCCTCAACCTACTACTGGGCTGCTTCT	6094
Dd	4955	AACTGCAATTTATTCATATCAGGATTTATCAATACCATATTTTGAAAAAGCCGTTCTG	5014	Qy	5060	AATGCAGGAGTTCATGGTAAATTTCTGACTTTTGGTTCATCAGTACAGTACGAAATGCT	5119
Qy	4829	-----	4828	Dd	6095	AATGCAGGAGTTCATGGTAAATTTCTGACTTTTGGTTCATCAGTACAGTACGAAATGCT	6154
Dd	5015	TAATGAAGGAGAAACTCAGGAGGAGTTTCCATAGGATGGCAAGATCCTGGTATCGGTC	5074	Qy	5120	GATACCGCATTTCTCAGTGTCTTGGGCTCTCTATCAGATTTGCCCACTAAAGCAAC	5179
Qy	4829	-----	4828	Dd	6155	GATACCGCATTTCTCAGTGTCTTGGGCTCTCTATCAGATTTGCCCACTAAAGCAAC	6214
Dd	5075	TGCGATTCCGACTCGTCCCAACATCAACCTATTAATTTCCCTCGTCAAAAATAAG	5134	Qy	5180	CGGAGGAGGAGTTCATGGTAAATTTCTGACTTTTGGTTCATCAGTACAGTACGAAATGCT	5239
Qy	4829	-----	4828	Dd	6215	CGGAGGAGGAGTTCATGGTAAATTTCTGACTTTTGGTTCATCAGTACAGTACGAAATGCT	6274
Dd	5135	GTATCAAGTGAGAAATCACCATGATGACGACCTGAATCCGGTGAGAAATGCAAAAGCTT	5194	Qy	5240	TGAGACTATCTCGGTTATGACAGCAAAATGTCCTTCTTGGGAGACAGTAAATGAAATGCC	5299
Qy	4829	-----	4828	Dd	6275	TGAGACTATCTCGGTTATGACAGCAAAATGTCCTTCTTGGGAGACAGTAAATGAAATGCC	6334
Dd	5195	ATGCATTTCTTCCAGACTTGTTCACAGGCCAGCCATACGCTCGTCAATCAAAATCACT	5254	Qy	5300	ACCAATAAGAAATCCTTGTATCAGAAACAACTTCTTGTTCGAACTTTTTCGGTGCC	5359
Qy	4829	-----	4828	Dd	6335	ACCAATAAGAAATCCTTGTATCAGAAACAACTTCTTGTTCGAACTTTTTCGGTGCC	6394
Dd	5255	CGCATCAACAAACCGTTATTCATTCTGATTGCGCTGAGCGACGAAATACGCGATC	5314	Qy	5360	TTGAACTTATAAATGTAGAGTGGATATGTCGGTAGGAAATGGAGCGGGGAAATGCTTACC	5419
Qy	4829	-----	4828	Dd	6395	TTGAACTTATAAATGTAGAGTGGATATGTCGGTAGGAAATGGAGCGGGGAAATGCTTACC	6454
Dd	5315	GCTGTTAAAGACAAATTAACAAAGGAATCGAATGCAACCGCGCGCAGAACACTGCCAG	5374	Qy	5420	TTCTGACCTTCAAGAGTATGAGGTTTGTAGTACTGATGCCAATTCAGTACGAGCAAC	5479
Qy	4829	-----	4828	Dd	6455	TTCTGACCTTCAAGAGTATGAGGTTTGTAGTACTGATGCCAATTCAGTACGAGCAAC	6514
Dd	5375	CGCATCAACAAATTTTACCTGAATCAGGATATCTTCTAATACCTCGAATGCTGTTTT	5434	Qy	5480	GTTGCTATTTCTGATTTCCGAATCCGAAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATT	5539
Qy	4829	-----	4828	Dd	6515	GTTGCTATTTCTGATTTCCGAATCCGAAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATT	6574
Dd	5435	CCCGGGATCGCAGTGGTGAATACCAACATCATCAGGAGTACGGATAAAATGTTGAT	5494	Qy	5540	GATCCAAAGCAGTGGCTTGTAAACTGACAAATAGTGTGCTGTTGTTTGGAGTCACTTT	5599
Qy	4829	-----	4828	Dd	6575	GATCCAAAGCAGTGGCTTGTAAACTGACAAATAGTGTGCTGTTGTTTGGAGTCACTTT	6634
Dd	5495	GGTCGAAGAGGCAATAATTCGTCAGCCAGTTTGTAGTCTGACCATCTCATCTGTAACATC	5554	Qy	5600	TGATGAATAAATCTAGTCTTTGATCTTAAATAATCTTGACGAGCAAGGCGGATAAATACC	5659



Db 6635 TGTATGAATAAATCTAGTCTTTGATCTAATAAATCTTTGACGAGCAAGCGATAAATAACC 6694  
Qy 5660 CAATCTAAATCTTTTAAACGTTAAAGGACAAGTATGTCGCGCTGATTAATAAACCCC 5719  
Db 6695 CAATCTAAATCTTTTAAACGTTAAAGGACAAGTATGTCGCGCTGATTAATAAACCCC 6754  
Qy 5720 AAATCAGCTCTAGTCTGATCCTCATCAACTTTGAGGGGCACTATCTTTGTTTATAGAAAT 5779  
Db 6755 AAAACAGCTCTAGTCTGATCCTCATCAACTTTGAGGGGCACTATCTTTGTTTATAGAAAT 6814  
Qy 5780 TTGGGAGATCGGATATCGAAGAAAAGTAGCTGATTTTAAACGTGAATTTATCTCAA 5839  
Db 6815 TTGGGAGATCGGATATCGAAGAAAAGTAGCTGATTTTAAACGTGAATTTATCTCAA 6874  
Qy 5840 GATCTCTCCCTCGCGCTTTCCGGTGATGACGGTGAAAACCTCTGACACATGCGAGTCCCG 5899  
Db 6875 GATCTCTCCCTCGCGCTTTCCGGTGATGACGGTGAAAACCTCTGACACATGCGAGTCCCG 6934  
Qy 5900 GAGCGGTCAAGCTTGTCTGTAAAGCGGATGCGGGAGACAGAACGCCGTGAGGGGCGG 5959  
Db 6935 GAGCGGTCAAGCTTGTCTGTAAAGCGGATGCGGGAGACAGAACGCCGTGAGGGGCGG 6994  
Qy 5960 TCAGCGGTGTTGGCGGTGTCGGGGCGCAGCCATGACCCAGTCAAGTACGATAGCGGA 6019  
Db 6995 TCAGCGGTGTTGGCGGTGTCGGGGCGCAGCCATGACCCAGTCAAGTACGATAGCGGA 7054  
Qy 6020 GTGTATCTGGCTTAACCTATCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC 6079  
Db 7055 GTGTATCTGGCTTAACCTATCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC 7114  
Qy 6080 GGTGTGAATACCGACAGATGCGTAGGAGAAAATACCGCATCAGCGGCTCTCCGCTT 6139  
Db 7115 GGTGTGAATACCGACAGATGCGTAGGAGAAAATACCGCATCAGCGGCTCTCCGCTT 7174  
Qy 6140 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCACT 6199  
Db 7175 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCACT 7234  
Qy 6200 CAAAGCGGTAAATACGGTTATTCACAGATCAGGGGATACCGAGGAAAAGACATGTGAG 6259  
Db 7235 CAAAGCGGTAAATACGGTTATTCACAGATCAGGGGATACCGAGGAAAAGACATGTGAG 7294  
Qy 6260 CAAAGGCGCAGAAAGCCAGGAACCGTAAAGGCGCGCTTCTCGCGCTTTTCCATA 6319  
Db 7295 CAAAGGCGCAGAAAGCCAGGAACCGTAAAGGCGCGCTTCTCGCGCTTTTCCATA 7354  
Qy 6320 GGCTCCGCCCTCGAGAGATCACAATAATCGACGCTCAAGTCAAGTGTGCGGAAC 6379  
Db 7355 GGCTCCGCCCTCGAGAGATCACAATAATCGACGCTCAAGTCAAGTGTGCGGAAC 7414  
Qy 6380 CGACGACTATAAGATACAGAGCTTTCCCGCTGGAAGCTCCCTCGCGCTCTCCTG 6439  
Db 7415 CGACGACTATAAGATACAGAGCTTTCCCGCTGGAAGCTCCCTCGCGCTCTCCTG 7474  
Qy 6440 TTCGACCTCCCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCG 6499  
Db 7475 TTCGACCTCCCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCG 7534  
Qy 6500 TTTCTCATAGCTCAGCTGATGATCTCAGTTGCTGAGTGGTGGCTTCCGCTCAAGCTGG 6559  
Db 7535 TTTCTCATAGCTCAGCTGATGATCTCAGTTGCTGAGTGGTGGCTTCCGCTCAAGCTGG 7594  
Qy 6560 GCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTC 6619  
Db 7595 GCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTC 7654  
Qy 6620 TTGAGTCAAACCGGTATAGACAGACTTATGCCACTGGCAGCAGCCACTGGTAAACAGGA 6679  
Db 7655 TTGAGTCAAACCGGTATAGACAGACTTATGCCACTGGCAGCAGCCACTGGTAAACAGGA 7714  
Qy 6680 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCTTAACCTACG 6739

Db 7715 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCTTAACCTACG 7774  
Qy 6740 GCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAA 6799  
Db 7775 GCTACACTAGAGGACAGTATTTGGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGGAA 7834  
Qy 6800 AAAGAGTTGGTAGCTCTTGTATCCGSCAAAACAAACACCGCTGGTAGCGGTGTTTTTGG 6859  
Db 7835 AAAGAGTTGGTAGCTCTTGTATCCGSCAAAACAAACACCGCTGGTAGCGGTGTTTTTGG 7894  
Qy 6860 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGAGATCTCAAGAAAGATCCTTTGATCTTTT 6919  
Db 7895 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGAGATCTCAAGAAAGATCCTTTGATCTTTT 7954  
Qy 6920 CTACGGGTCTGACGCTCAGTGGAAACGAAACTCAAGTAAAGGATTTTGGTCTATGAGAT 6979  
Db 7955 CTACGGGTCTGACGCTCAGTGGAAACGAAACTCAAGTAAAGGATTTTGGTCTATGAGAT 8014  
Qy 6980 TATCAAAAGAGATCTTCACTAGATCTCTTTTAAATTTAAATGAAGTTTAAATCAATCT 7039  
Db 8015 TATCAAAAGAGATCTTCACTAGATCTCTTTTAAATTTAAATGAAGTTTAAATCAATCT 8074  
Qy 7040 AAAGTATATATGAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAAATCAGTCAGGACCTA 7099  
Db 8075 AAAGTATATATGAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAAATCAGTCAGGACCTA 8134  
Qy 7100 TCTCAGCATCTGCTCTATTTGTTTCATCCATAGTTGCTGATCTCCCTGCTGTAGATAA 7159  
Db 8135 TCTCAGCATCTGCTCTATTTGTTTCATCCATAGTTGCTGATCTCCCTGCTGTAGATAA 8194  
Qy 7160 CTACCATACGGAGGGCTTACCATCTGGCCCGCAGTGTGCAATGATACCGGAGACCCAC 7219  
Db 8195 CTACCATACGGAGGGCTTACCATCTGGCCCGCAGTGTGCAATGATACCGGAGACCCAC 8254  
Qy 7220 GTCACCGGCTCCAGATTATCAGCAATAAACACGACGACCGGAGGCGCCAGCGCAGAA 7279  
Db 8255 GTCACCGGCTCCAGATTATCAGCAATAAACACGACGACCGGAGGCGCCAGCGCAGAA 8314  
Qy 7280 GTGGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAAAGTAGAG 7339  
Db 8315 GTGGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAAAGTAGAG 8374  
Qy 7340 TAAAGTGTGCGCACTAATAGTTGCGCAACGTTGTTGCCATTGCTCGACGGCTCGTGG 7399  
Db 8375 TAAAGTGTGCGCACTAATAGTTGCGCAACGTTGTTGCCATTGCTCGACGGCTCGTGG 8434  
Qy 7400 TGTACGCTCCGCTGCTGATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGGAG 7459  
Db 8435 TGTACGCTCCGCTGCTGATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGGAG 8494  
Qy 7460 TTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTGG 7519  
Db 8495 TTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTGG 8554  
Qy 7520 TCAGAAAGTGTGCGCGCAGTGTATCACTCATGTTATGCGACGCTGATTAATCTC 7579  
Db 8555 TCAGAAAGTGTGCGCGCAGTGTATCACTCATGTTATGCGACGCTGATTAATCTC 8614  
Qy 7580 TTTACTGTATGCGCATCCGCTAAGATGCTTTTCTGTGACCTGGTGAAGTCTCAACCAAGTCA 7639  
Db 8615 TTTACTGTATGCGCATCCGCTAAGATGCTTTTCTGTGACCTGGTGAAGTCTCAACCAAGTCA 8674  
Qy 7640 TCTGAGATATGTTATGCGGCGACCGAGTTGCTTGTCCCGCGCTCAACACGGGATATA 7699  
Db 8675 TCTGAGATATGTTATGCGGCGACCGAGTTGCTTGTCCCGCGCTCAACACGGGATATA 8734  
Qy 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAGCGTTCTTCGGGGCGAA 7759  
Db 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAGCGTTCTTCGGGGCGAA 8794  
Qy 7760 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 7819  
Db 8795 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 8854

```
QY 7820 ACTGATCTTACGATCTTTTACTTTTCCACGAGTTTCTGGTGAGGAAAACAGGAAGGC 7879
Db 8855 ACTGATCTTACGATCTTTTACTTTTCCACGAGTTTCTGGTGAGGAAAACAGGAAGGC 8914
QY 7880 AAAATGCGCGAAAAAGGGAATAAGGCGACACGGAATGTTGAATCTATCTCTTCC 7939
Db 8915 AAAATGCGCGAAAAAGGGAATAAGGCGACACGGAATGTTGAATCTATCTCTTCC 8974
QY 7940 TTTTTCATATATTGTAAGCAATTTATCAGGGTTTATGTCATCAGCGGATACATATTG 7999
Db 8975 TTTTTCATATATTGTAAGCAATTTATCAGGGTTTATGTCATCAGCGGATACATATTG 9034
QY 8000 AATGATTTTGAAGAAATAAACAATAGGTTTCCGCGACATTTCCCGAAAAGTGCCAC 8059
Db 9035 AATGATTTTGAAGAAATAAACAATAGGTTTCCGCGACATTTCCCGAAAAGTGCCAC 9094
QY 8060 CTGACGCTAAGAAACCAATTTATCATGACATTAACCTATAAATAAGGCGTATACGA 8119
Db 9095 CTGACGCTAAGAAACCAATTTATCATGACATTAACCTATAAATAAGGCGTATACGA 9154
QY 8120 GGCCCTTTGCTCTCAAGAAATTAATCTCATGTTTGACAGCTTATCATGTAAGCTGAC 8179
Db 9155 GGCCCTTTGCTCTCAAGAAATTAATCTCATGTTTGACAGCTTATCATGTAAGCTGAC 9214
QY 8180 TCATGTTGCTATTGTGAATAAGACGACATCGGAAACACTGAAAAATAACAGTTATTATT 8239
Db 9215 TCATGTTGCTATTGTGAATAAGACGACATCGGAAACACTGAAAAATAACAGTTATTATT 9274
QY 8240 CG 8241
Db 9275 CG 9276
```

```
RESULT 11
ID AAD47231 standard; DNA; 9276 BP.
XX AC AAD47231;
XX DT 24-FEB-2003 (first entry)
XX DE PPC9K vector DNA containing human MTSP9 sequence.
KW Human; type II membrane-type serine protease 9; tumour; transgenic;
KW type II transmembrane serine protease; enzyme; gene therapy; MTSP9;
KW neoplastic disease; transgenic animal; TTSP; ds.
XX OS Homo sapiens.
XX OS Pichia pastoris.
XX OS Unidentified.
XX OS Chimeric.
XX PH Key
XX FT Promoter
XX FT Location/Qualifiers
XX FT 1..948
XX FT /tag= a
XX FT /note= "AOX1 promoter fragment"
XX FT primer_bind
XX FT 855..875
XX FT /tag= b
XX FT /bound_moiety= "AOX1 primer"
XX FT misc_feature
XX FT 949..1218
XX FT /tag= c
XX FT /note= "Alpha-factor secretory signal sequence"
XX FT primer_bind
XX FT 1152..1172
XX FT /tag= d
XX FT /bound_moiety= "Alpha-factor primer"
XX FT misc_feature
XX FT 1192..1241
XX FT /tag= e
XX FT /note= "Multiple cloning site"
XX FT misc_signal
XX FT 1253..1586
XX FT /tag= g
XX FT /note= "AOX1 transcription termination region"
XX FT primer_bind
XX FT 1327..1347
```

```
FT /tag= f
FT /bound_moiety= "AOX1 primer"
FT complement(1980..4514)
FT /tag= h
FT /note= "HIS4 ORF"
FT misc_feature
FT complement(4928..5743)
FT /tag= i
FT /note= "Kanamycin resistance gene"
FT misc_feature
FT 6122..6879
FT /tag= j
FT /note= "AOX1 fragment"
FT misc_feature
FT complement(7288..7961)
FT /tag= k
FT /note= "ColE1 origin"
FT misc_feature
FT complement(8106..8966)
FT /tag= l
FT /note= "Ampicillin resistance gene"
XX WO200277267-A2.
XX PN 03-OCT-2002.
XX 27-MAR-2002; 2002WO-US009611.
XX 27-MAR-2001; 2001US-0279228P.
XX 15-MAY-2001; 2001US-0291501P.
XX (CORV-) CORVAS INT INC.
XX Madison EL, Ong EO;
XX WPI; 2003-018940/01.
XX New substantially purified single or two-chain type II membrane-type
XX serine protease 9 (MTSP9) polypeptide, useful for monitoring tumor
XX progression, inhibiting tumor initiation, or treating a malignant or pre-
XX malignant condition.
XX Example 2; Page 191-193; 199pp; English.
XX The invention relates to type II membrane-type serine protease 9 (MTSP9)
XX polypeptides and polynucleotides. MTSP9 belongs to type II transmembrane
XX serine protease (TSP) family. Sequences of the invention and their
XX antibodies are useful for diagnosing, treating or preventing neoplastic
XX disease in mammals. They are useful for monitoring tumour progression,
XX inhibiting tumour initiation, growth or progression or treating malignant
XX or pre-malignant conditions. Transgenic animals of the invention are
XX useful in animal models of tumour initiation, growth and/or progression
XX models. The invention is also useful in gene therapy. The present
XX sequence is PPC9K vector DNA containing human MTSP9 sequence
XX SQ Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;
```

```
Query Match 79.6%; Score 6561.6; DB 7; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;
QY 1 AGATCTAACATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
Db 1 AGATCTAACATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
QY 61 GTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACACTAGCAGCAGCCGT 120
Db 61 GTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACACTAGCAGCAGCCGT 120
QY 121 TGCAAAACGCGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGGCATCGAAAACC 180
Db 121 TGCAAAACGCGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGGCATCGAAAACC 180
QY 181 AGCCAGTATTGGGCTTGATTTGGAGCTCGCTCATTTCCATTCCTTCTATTAGGCTACTA 240
Db 181 AGCCAGTATTGGGCTTGATTTGGAGCTCGCTCATTTCCATTCCTTCTATTAGGCTACTA 240
```

QY 241 ACACATGACCTTATTAGCCTGCTATCTCTGGCCCTCGCGAGGTTTCATGTTTGTGTTA 300  
Db 241 ACACATGACCTTATTAGCCTGCTATCTCTGGCCCTCGCGAGGTTTCATGTTTGTGTTA 300  
QY 301 TTTCCGAATGCAACAGCTCGCATTAACACCCGAAACATCACTCCAGATGAGGCTTTCTG 360  
Db 301 TTTCCGAATGCAACAGCTCGCATTAACACCCGAAACATCACTCCAGATGAGGCTTTCTG 360  
QY 361 AGTGTGGGTCAAAATAGTTTCATGTTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420  
Db 361 AGTGTGGGTCAAAATAGTTTCATGTTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420  
QY 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAATAGTTTGGTTG 480  
Db 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAATAGTTTGGTTG 480  
QY 481 TTGAATGCTAACGCCAGTTGTTCAAAAAGAACTTCCAAAAGTCCGCCATACCGTTTGT 540  
Db 481 TTGAATGCTAACGCCAGTTGTTCAAAAAGAACTTCCAAAAGTCCGCCATACCGTTTGT 540  
QY 541 CTCTGTTGGTATTGATTGACGAATGCTCAAAAATTAATCTCAATTAATGCTTAGCGCAGTCT 600  
Db 541 CTCTGTTGGTATTGATTGACGAATGCTCAAAAATTAATCTCAATTAATGCTTAGCGCAGTCT 600  
QY 601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGGAAACCAAATGGGGAAAACCCCGCT 660  
Db 601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGGAAACCAAATGGGGAAAACCCCGCT 660  
QY 661 TTTTGGATGATTGCAATGCTCCACATGCTATGCTTCCAGATTCCTGTTGGGAATACT 720  
Db 661 TTTTGGATGATTGCAATGCTCCACATGCTATGCTTCCAGATTCCTGTTGGGAATACT 720  
QY 721 GCTGATAGCCTAACGTTTCATGATCAAAATTAATGCTTCTAACCCCTACTTGACAGCAAT 780  
Db 721 GCTGATAGCCTAACGTTTCATGATCAAAATTAATGCTTCTAACCCCTACTTGACAGCAAT 780  
QY 781 ATATAACAGAAAGTGCCTGCTTAAACCTTTTATCATCATTAATAGCTT 840  
Db 781 ATATAACAGAAAGTGCCTGCTTAAACCTTTTATCATCATTAATAGCTT 840  
QY 841 ACTTTCATTAATGCGACTGTTCCAAATGACAGCTTTTGATTTAAACGACTTTTAAACGA 900  
Db 841 ACTTTCATTAATGCGACTGTTCCAAATGACAGCTTTTGATTTAAACGACTTTTAAACGA 900  
QY 901 CAACCTTGAGAGATCAAAAACAATTAATTCGAAGGATCCAAACGATGAGATTTCCT 960  
Db 901 CAACCTTGAGAGATCAAAAACAATTAATTCGAAGGATCCAAACGATGAGATTTCCT 960  
QY 961 TCAATTTTACTGAGTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACT 1020  
Db 961 TCAATTTTACTGAGTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACT 1020  
QY 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTCAATCGGTTACTCAGATTTA 1080  
Db 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTCAATCGGTTACTCAGATTTA 1080  
QY 1081 GAAGGGATTCGATGTTGCTGTTTCCCATTTTCCAAAGCAGCAATAACGGGTTATTG 1140  
Db 1081 GAAGGGATTCGATGTTGCTGTTTCCCATTTTCCAAAGCAGCAATAACGGGTTATTG 1140  
QY 1141 TTTATAATACTACTATTGCGAGATTGCTGCTAAAGAAGAGGGTATCTCTCGAGAAA 1200  
Db 1141 TTTATAATACTACTATTGCGAGATTGCTGCTAAAGAAGAGGGTATCTCTCGAGAAA 1200  
QY 1201 AGAGAGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGAGAAACATGATCATG 1260  
Db 1201 AGAGAGCTGAAGCTA----- 1217  
QY 1261 GCCCCCCGGAGAGGATAAATTGTGGCTTCCCGGTGTCCCGCCAGCAGTGCAGGAG 1320  
Db 1218 ----- 1217  
QY 1321 AGAGGTGCTGTTTGTATGACAGTGTCCGGGGATTCGGTGTCTCCACCCCATGGCC 1380

Db 1218 ----- 1217  
QY 1381 ATCGAGAACACTCAAGAAGAAGATGTCCTTTCTAACTAGTGGCGTAGAATTCCTTAGGG 1440  
Db 1218 -----CGTAGAATTCCTTAGGG 1234  
QY 1441 CGGCGCGAATTAATTCGCCCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTTACG 1500  
Db 1235 CGGCGCGAATTAATTCGCCCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTTACG 1234  
QY 1501 AGAAGACCGGTCTTCGTAGATTCTTAATCAAGAGATGTCAGAATGCCATTTCCTGAGAG 1560  
Db 1295 AGAAGACCGGTCTTCGTAGATTCTTAATCAAGAGATGTCAGAATGCCATTTCCTGAGAG 1354  
QY 1561 ATGCAAGGCTTCATTTTGTGATACCTTTTATTTGTAACCTATATAGTATAGATTTTTTTT 1620  
Db 1355 ATGCAAGGCTTCATTTTGTGATACCTTTTATTTGTAACCTATATAGTATAGATTTTTTTT 1414  
QY 1621 GTCAATTTTGTCTTCGTACGAGCTTGCTCTGATCAGCCTATCTCGAGCTGATGAA 1680  
Db 1415 GTCAATTTTGTCTTCGTACGAGCTTGCTCTGATCAGCCTATCTCGAGCTGATGAA 1474  
QY 1681 TATCTGTGTAGGGTTTGGGAAATCATTTCGAGTTTGATGTTTCTTGTTGATTTTCCC 1740  
Db 1475 TATCTGTGTAGGGTTTGGGAAATCATTTCGAGTTTGATGTTTCTTGTTGATTTCCC 1534  
QY 1741 ACTCCTCTTCAGAGTACAGAAAGATTAAAGTGAGAAAGTTTCGTTTGTGCAAGCTTATCGATAA 1800  
Db 1535 ACTCCTCTTCAGAGTACAGAAAGATTAAAGTGAGAAAGTTTCGTTTGTGCAAGCTTATCGATAA 1594  
QY 1801 GCTTTAATGCGGTAGTTTATCAGAGTTAAATGCTTAACGAGTACAGCAGCGTATGAA 1860  
Db 1595 GCTTTAATGCGGTAGTTTATCAGAGTTAAATGCTTAACGAGTACAGCAGCGTATGAA 1654  
QY 1861 ATCTAAACATGCGTCACTGCTCCTCGGCAACCGTCAACCTCGATGCTGTAGGCATAGG 1920  
Db 1655 ATCTAAACATGCGTCACTGCTCCTCGGCAACCGTCAACCTCGATGCTGTAGGCATAGG 1714  
QY 1921 CTTGTTTATGCGGTAGTGTGCGGGCTCTTCGCGGATATCGTCCATTCGACAGCATCGC 1980  
Db 1715 CTTGTTTATGCGGTAGTGTGCGGGCTCTTCGCGGATATCGTCCATTCGACAGCATCGC 1774  
QY 1981 CAGTCACATATGCGGTGCTAGCGCTATATGCGTGTGATGCAATTTCTATCGCACCCGT 2040  
Db 1775 CAGTCACATATGCGGTGCTAGCGCTATATGCGTGTGATGCAATTTCTATCGCACCCGT 1834  
QY 2041 TCTCGGAGCACTGTCCGACCGCTTTGCGCGCCCGCCAGTCTGCTGCTTCGCTACTTGG 2100  
Db 1835 TCTCGGAGCACTGTCCGACCGCTTTGCGCGCCCGCCAGTCTGCTGCTTCGCTACTTGG 1894  
QY 2101 AGCCATATCGACTACCGGATCATGGGACCAACCCGCTCTGTGGATCTATCGAATCTA 2160  
Db 1895 AGCCATATCGACTACCGGATCATGGGACCAACCCGCTCTGTGGATCTATCGAATCTA 1954  
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2220  
Db 1955 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2014  
QY 2221 CAGCATTCGGGTAGCATCTAGACTTCAACAGAGCCAGATCCATCATCTCTTGGCCAA 2280  
Db 2015 CAGCATTCGGGTAGCATCTAGACTTCAACAGAGCCAGATCCATCATCTCTTGGCCAA 2074  
QY 2281 TATGTTTCAGTCCCTCAGGAGTTAGTCTTGTGAAGTGAATGAACTTCTGGAAGTTGCAG 2340  
Db 2075 TATGTTTCAGTCCCTCAGGAGTTAGTCTTGTGAAGTGAATGAACTTCTGGAAGTTGCAG 2134  
QY 2341 TGTTAATCCCGTGTATTGACGGGCATATCCGTAGCTTGGCAAGTGTGTTGTACCGG 2400  
Db 2135 TGTTAATCCCGTGTATTGACGGGCATATCCGTAGCTTGGCAAGTGTGTTGTACCGG 2194  
QY 2401 AGGAGTAAATCTCCAACTCTCTCGAGAGTAGGCACCAACAAACACAGATCCAGCGTGT 2460

2195	AGGAGTATCTCCACAACTCTCTGGAGAGTAGGCCAACCAACACACAGATCCAGCGTGTT	2254
2461	GTACTTGATCAACATTAAGAAAGACATCTCTCGATTTCAGGATCAAGTGTTCAGAGCGT	2520
2255	GTACTTGATCAACATTAAGAAAGACATCTCTCGATTTCAGGATCAAGTGTTCAGAGCGT	2314
2521	ACTGATTGGACATTTCCAAAGCGCTGCTCGTAGGTTGCCAACCGATAGGGTTGTAGAGTGTG	2580
2315	ACTGATTGGACATTTCCAAAGCGCTGCTCGTAGGTTGCCAACCGATAGGGTTGTAGAGTGTG	2374
2581	CAATACACTTGGGTACAATTTCAACCCCTTGGGCAACTGCACACTGTTGTTGTGAACAGCAT	2640
2375	CAATACACTTGGGTACAATTTCAACCCCTTGGGCAACTGCACACTGTTGTTGTGAACAGCAT	2434
2641	CTTCAATTCTGSCAAGCTCCTTGTCTCATATCGACAGCCAAACAGAAATCACTCGGGAAT	2700
2435	CTTCAATTCTGSCAAGCTCCTTGTCTCATATCGACAGCCAAACAGAAATCACTCGGGAAT	2494
2701	CAATACCATGTTTCAGCTTCGAGACAGAAGGCTGTGAGGCAACGAAATCTCGATCAGCGTATT	2760
2495	CAATACCATGTTTCAGCTTCGAGACAGAAGGCTGTGAGGCAACGAAATCTCGATCAGCGTATT	2554
2761	TATCAGCAATAACTAGAACTTTCAGAGAGGCCAGCAGGCATGTCAATACTACACAGGCTG	2820
2555	TATCAGCAATAACTAGAACTTTCAGAGAGGCCAGCAGGCATGTCAATACTACACAGGCTG	2614
2821	ATGTGTCAATTTTGAACCAATCATCTTGGCAGCAGTAACGAACTGGTTTCTCTGGACCAATA	2880
2615	ATGTGTCAATTTTGAACCAATCATCTTGGCAGCAGTAACGAACTGGTTTCTCTGGACCAATA	2674
2881	TTTTTGTCACTTAGAAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTGGCGC	2940
2675	TTTTTGTCACTTAGAAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTGGCGC	2734
2941	CTCCTGCTAGCAGATACACTTAGACCAACCTTGTGGGCAACGTAGATGACTTCTGGGG	3000
2735	CTCCTGCTAGCAGATACACTTAGACCAACCTTGTGGGCAACGTAGATGACTTCTGGGG	2794
3001	TAAAGGTACCATCCTTCTTAGTGTGAGATGCAAAAACAATTTCTTTGCAACCAAGCACTT	3060
2795	TAAAGGTACCATCCTTCTTAGTGTGAGATGCAAAAACAATTTCTTTGCAACCAAGCACTT	2854
3061	TGCGAGGAACACCCAGCATCAGGAGAGTGGAAGGCAGAAATGCGGTTCCACACAGGAATAT	3120
2855	TGCGAGGAACACCCAGCATCAGGAGAGTGGAAGGCAGAAATGCGGTTCCACACAGGAATAT	2914
3121	AGAGGGCAACTTTCTCAATAGGCTTTGCCAAAACGAGAGCAGACTACACAGGCAAGTCT	3180
2915	AGAGGGCAACTTTCTCAATAGGCTTTGCCAAAACGAGAGCAGACTACACAGGCAAGTCT	2974
3181	CAACTTCACAGCTCCGTTAGTTGAGCTTTCATGGAATTTCCGTAGCTTATCTATAGAGA	3240
2975	CAACTTCACACGCTCCGTTAGTTGAGCTTTCATGGAATTTCCGTAGCTTATCTATAGAGA	3034
3241	GATCAATGGCTCTCTTAACCGTTATCTGGCAATTTGCATAAGTTCTCTGGGAAAAGGAGCTT	3300
3035	GATCAATGGCTCTCTTAACCGTTATCTGGCAATTTGCATAAGTTCTCTGGGAAAAGGAGCTT	3094
3301	CTAACACAGGTGCTTCAAAAGCGACTCCATCAAACTTGGCAGTAGTTCTTAAAGGGCTT	3360
3095	CTAACACAGGTGCTTCAAAAGCGACTCCATCAAACTTGGCAGTAGTTCTTAAAGGGCTT	3154
3361	TGTCACCAATTTTGACGAACATTTGTCGACAAATGTTTTGACTAATTCCTAATCTGTTCCG	3420
3155	TGTCACCAATTTTGACGAACATTTGTCGACAAATGTTTTGACTAATTCCTAATCTGTTCCG	3214
3421	TTTTTCTGGATAGGACGACGAAGGCATCTTCAATTTCTTGTGAGGAGGCCCTTAGAACGT	3480
3215	TTTTTCTGGATAGGACGACGAAGGCATCTTCAATTTCTTGTGAGGAGGCCCTTAGAACGT	3274
3481	CAATTTTGCACAAATTCATACGACTTTCAGAGGCACTCTTTTAGGTTTCGATTTCTCTT	3540
3275	CAATTTTGCACAAATTCATACGACTTTCAGAGGCACTCTTTTAGGTTTCGATTTCTCTT	3334

QY	3541	TAGGTTGTTCCCTTGGTGATCTCGGCTTGGCATCTCCTTCTCTAGTGACCTTTAGGG	3600
DB	3335	TAGGTTGTTCCCTTGGTGATCTCGGCTTGGCATCTCCTTCTAGTGACCTTTAGGG	3394
QY	3601	ACTTCATATCCAGGTTTCTCTCCACTCGTCCAAAGTCCACCGTACTTGGCACATCTAA	3660
DB	3395	ACTTCATATCCAGGTTTCTCTCCACTCGTCCAAAGTCCACCGTACTTGGCACATCTAA	3454
QY	3661	CTAATCAAAATAAAATAAGTCAGCACATTCCCAGGCTATATCTTCTTGGATTAGCTT	3720
DB	3455	CTAATCAAAATAAAATAAGTCAGCACATTCCCAGGCTATATCTTCTTGGATTAGCTT	3514
QY	3721	CTGCAAGTTCATCAGGTTCTCTCGCTTAATTTTAGCGTTCAACAAAACCTTCGTCCAAATA	3780
DB	3515	CTGCAAGTTCATCAGGTTCTCTCGCTTAATTTTAGCGTTCAACAAAACCTTCGTCCAAATA	3574
QY	3781	ACCGTTTGGTATAAGAAACCTTCTGGAGCATGCTCTTACGATCCCAAGGTGGCTTCCA	3840
DB	3575	ACCGTTTGGTATAAGAAACCTTCTGGAGCATGCTCTTACGATCCCAAGGTGGCTTCCA	3634
QY	3841	TGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGTTTCCAAAGTCAGAAACCAA	3900
DB	3635	TGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGTTTCCAAAGTCAGAAACCAA	3694
QY	3901	CACCTGTTTGTTCAAACCAAAATTTCAAGCAGGCTCTCATCACAATCCAAATTCGATACCCA	3960
DB	3695	CACCTGTTTGTTCAAACCAAAATTTCAAGCAGGCTCTCATCACAATTCGATACCCA	3754
QY	3961	GCAACTTTTGAGTTGTCOCAGATGTTAGCAGCTTTATACCAAAAACGTCAGACGAGATT	4020
DB	3755	GCAACTTTTGAGTTGTCOCAGATGTTAGCAGCTTTATACCAAAAACGTCAGACGAGATT	3814
QY	4021	GGTAGACTCCAGTTTGTGCTTATAGCTTCGGGAATAGACTTTTGGACGAGTACACCA	4080
DB	3815	GGTAGACTCCAGTTTGTGCTTATAGCTTCGGGAATAGACTTTTGGACGAGTACACCA	3874
QY	4081	GGGCCAACGAGTAAATTAGAGAGTTCAGCCACCAAAAGTAGTAATAGACCATCGGGCGGT	4140
DB	3875	GGGCCAACGAGTAAATTAGAGAGTTCAGCCACCAAAAGTAGTAATAGACCATCGGGCGGT	3934
QY	4141	CAGTAGTCAAAAGCCGCAAAAATTTCACTGACGAGGACCTTTTGGACATCTTCAGAAA	4200
DB	3935	CAGTAGTCAAAAGCCGCAAAAATTTCACTGACGAGGACCTTTTGGACATCTTCAGAAA	3994
QY	4201	GTTTCGATTTCAGTAGTCAATTTGCCAGGATCAATAATGGGGATTATACCAAGAACACAG	4260
DB	3995	GTTTCGATTTCAGTAGTCAATTTGCCAGGATCAATAATGGGGATTATACCAAGAACACAG	4054
QY	4261	TGGAAGTCAATCTACCAACTTTTGGCGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC	4320
DB	4055	TGGAAGTCAATCTACCAACTTTTGGCGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC	4114
QY	4321	CATTAGTGAACATTTTCAAATTCGCCAGTGGAGAAAGAAAAGGCACAGCGCATACTAGCAT	4380
DB	4115	CATTAGTGAACATTTTCAAATTCGCCAGTGGAGAAAGAAAAGGCACAGCGCATACTAGCAT	4174
QY	4381	TAGCGGCAAGGATGCAACTTTATCAACAGGGTCTATAGATAAACCTTAGCGCTGGGA	4440
DB	4175	TAGCGGCAAGGATGCAACTTTATCAACAGGGTCTATAGATAAACCTTAGCGCTGGGA	4234
QY	4441	TCATCCTTTGGACAACCTTTTCTGCCAAATCTTAGGTCCAAAATCACCTTCATTGATACCAT	4500
DB	4235	TCATCCTTTGGACAACCTTTTCTGCCAAATCTTAGGTCCAAAATCACCTTCATTGATACCAT	4294
QY	4501	TATTGTAACTTGAGCAAGTTGCGATAGCTTCCTCAAAATGGTCTCTGTAAACGATG	4560
DB	4295	TATTGTAACTTGAGCAAGTTGCGATAGCTTCCTCAAAATGGTCTCTGTAAACGATG	4354
QY	4561	ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTGAGTGAACCTTCAGGTTGTGCA	4620
DB	4355	ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTGAGTGAACCTTCAGGTTGTGCA	4414

QY 4621 GCTGTCAGCAGCATAGGAAACACGGCTTTTCTACCAAACTCAAGGAATTTCAAACCT 4680  
Db 4415 GCTGTCAGCAGCATAGGAAACACGGCTTTTCTACCAAACTCAAGGAATTTCAAACCT 4474  
QY 4681 CTGCAACACTTCCGTATCCAGTAGCAGGGAATGTCATCTGAAGTCGGACAGTCAG 4740  
Db 4475 CTGCAACACTTCCGTATCCAGTAGCAGGGAATGTCATCTGAAGTCGGACAGTCAG 4534  
QY 4741 TGTAGTCCTTGAGAAATCTGAAGCCGTATTTTATTATCATGAGTCAGTCATCAGGAGA 4800  
Db 4535 TGTAGTCCTTGAGAAATCTGAAGCCGTATTTTATTATCATGAGTCAGTCATCAGGAGA 4594  
QY 4801 TCCTCTACGCCGAGCGCATCTGGCCGA----- 4828  
Db 4595 TCCTCTACGCCGAGCGCATCTGGCCGA----- 4654  
QY 4829 ----- 4828  
Db 4655 CTCGTGAGAGAGGTGTTGCTGACTCATACAGGCGCTGNAATCGCCCATCATCCAGCCAGA 4714  
QY 4829 ----- 4828  
Db 4715 AAGTAGGAGGCCAGGTTGTATGAGAGCTTTGTTGTAGGTGACCAAGTTGGTATTGGA 4774  
QY 4829 ----- 4828  
Db 4775 ACTTTTGCTTTGCCACGAAACGGTCTCGGTTGTCCGGAAGATGCGTGATCTGATCTCTCA 4834  
QY 4829 ----- 4828  
Db 4835 ACTCAGCAAAAGTTCGATTTATTCAAACAAGCCGCTCCGCTCAAGTCAGCGTAATGCT 4894  
QY 4829 ----- 4828  
Db 4895 CTGCCAGTGTACAAACCAATTAACCAATCTGATTAGAAAACTCATCGAGCATCAAAATG 4954  
QY 4829 ----- 4828  
Db 4955 AAAGTCAATTTATTATCATCAGGATTTATCAATACCATATTTTGTGAAAGCCGTTTCG 5014  
QY 4829 ----- 4828  
Db 5015 TAATGAAGGAGAAACTCACCGAGGAGTTCCATPAGGATGGCAAGATCCCTGGTATCGTCT 5074  
QY 4829 ----- 4828  
Db 5075 TGCAGTCCGACTCGTCCACATCAATACAACTTAAATTTCCCTCTCGTCAAAATAG 5134  
QY 4829 ----- 4828  
Db 5135 GTTATCAAGTAGAAATCACCATCAGTACGAGTGAATCCGGTGAAGATGGCAAAAGCTT 5194  
QY 4829 ----- 4828  
Db 5195 ATGCATTTCTTCCAGACTTGTTCACAGGCCAGCCATTACGCTCGTCATCAAAATCACT 5254  
QY 4829 ----- 4828  
Db 5255 GGCATCAACCAAAACCGTTATTCTATTCTGTGATTGCGCTGAGCGAGACGAATACGGATC 5314  
QY 4829 ----- 4828  
Db 5315 GCTGTTAAAGGACAAATTAACAAAGGAATCGAATGCAACCGGCGAGGAACACTGCCAG 5374  
QY 4829 ----- 4828  
Db 5375 GGCATCAACAAATTTTTCACCTGAAATCAGGATATTCTTCTAATACCTGGAATGCTGTTT 5434  
QY 4829 ----- 4828  
Db 5435 CCCGGGATCGAGTGTGATTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGAT 5494  
QY 4829 ----- 4828

Db 5495 GGTGGAAGAGGCATAAATTCGTGAGCCAGTTTAGTCTGACCATCTCATCTGTAACATC 5554  
QY 4829 ----- 4828  
Db 5555 ATTGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGCTTCCCAT 5614  
QY 4829 ----- 4828  
Db 5615 CAATCGATAGATTGTGCAACCTGATTGCCGACATTTATCGGAGCCCATTTATACCCATA 5674  
QY 4829 ----- 4828  
Db 5675 TAAATCAGATCCATGTTTGGAAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5734  
QY 4829 ----- 4828  
Db 5735 ATGGCTCATACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCTAGA 5794  
QY 4829 ----- 4828  
Db 5795 TGATATATTTTATCTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTC 5854  
QY 4829 ----- 4828  
Db 5855 CCCCCCCCCCTGACAGTTCGCAATCACCGCGCCACAGGTGCGGTGCTGCGGCTATAT 5914  
QY 4880 CGCCGACATCACCGATGCGGAAGATCGGCTCGCACTTCGGGCTCATGAGCGTGTGTTT 4939  
Db 5915 CGCCGACATCACCGATGCGGAAGATCGGCTCGCACTTCGGGCTCATGAGCGTGTGTTT 5974  
QY 4940 CGGCTGTGATGTTGGCAGGCCCGCTGGCGGGGACTGTTGGCGGCACTCTCCTTGCA 4999  
Db 5975 CGGCTGTGATGTTGGCAGGCCCGCTGGCGGGGACTGTTGGCGGCACTCTCCTTGCA 6034  
QY 5000 TGCACCATTCCTTGGCGCGGCTGCTCAACGGGCTCAACCTACTACTGCGGCTCTTCTCT 5059  
Db 6035 TGCACCATTCCTTGGCGCGGCTGCTCAACGGGCTCAACCTACTACTGCGGCTCTTCTCT 6094  
QY 5060 TATCGAGAGTCGATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAATGCT 5119  
Db 6095 AATGAGAGTCGATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAATGCT 6154  
QY 5120 GATACCCCATTCCTTCACTGCTCTTGAAGTCTCTTATCAGATTATGCCCACTAAAGCAAC 5179  
Db 6155 GATACCCCATTCCTTCACTGCTCTTGAAGTCTCTTATCAGATTATGCCCACTAAAGCAAC 6214  
QY 5180 CGGAGGAGAGATTCATGTTAAATTTCTGACTTTTGGTTCATCAGTAGCTGAACTG 5239  
Db 6215 CGGAGGAGAGATTCATGTTAAATTTCTGACTTTTGGTTCATCAGTAGCTGAACTG 6274  
QY 5240 TGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTTTGGAGACAGTAATGAAGTCCC 5299  
Db 6275 TGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTTTGGAGACAGTAATGAAGTCCC 6334  
QY 5300 ACCAATAAGAAATCTTGTGTTATCAGGAACAACTCTTGTGTTTGAACCTTTTTCGGTGCC 5359  
Db 6335 ACCAATAAGAAATCTTGTGTTATCAGGAACAACTCTTGTGTTTGAACCTTTTTCGGTGCC 6394  
QY 5360 TTGAACATAAATCTAGAGTGGATATGTCGGGTAGGAATGCGAGCGGCAAACTCTTACC 5419  
Db 6395 TTGAACATAAATCTAGAGTGGATATGTCGGGTAGGAATGCGAGCGGCAAACTCTTACC 6454  
QY 5420 TTCTGGAACCTTCAAGAGTATGTAGGTTTGTATGATCTGATGCAACTCTCAGTGCAAC 5479  
Db 6455 TTCTGGAACCTTCAAGAGTATGTAGGTTTGTATGATCTGATGCAACTCTCAGTGCAAC 6514  
QY 5480 GTTGCTATTTCTGTTCAACACCATTCGGAATCCAGAGAAATCAAGTCTGTTGCTACTATT 5539  
Db 6515 GTTGCTATTTCTGTTCAACACCATTCGGAATCCAGAGAAATCAAGTCTGTTGCTACTATT 6574  
QY 5540 GATCCAGCCAGTCGGTCTTTGAAACTGCAATAGTGTGCTCGTGTGTTTGAAGTCACTTT 5599

Db 6575 GATCAAGCCAGTCGGCTTTGAAACTGACAATAGTGTGCTCGTGTGTTTGGAGTCATCTT 6634  
Qy 5600 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTTGAGGCAAGGCGGATAAATACC 5659  
Db 6635 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTTGAGGCAAGGCGGATAAATACC 6694  
Qy 5660 CAAATCTAAATCTCTTTTAAAGCTTAAAGGCAAGATATGTCGCTGTATTAACCCC 5719  
Db 6695 CAAATCTAAATCTCTTTTAAAGCTTAAAGGCAAGATATGTCGCTGTATTAACCCC 6754  
Qy 5720 AAATCAGCTCGTGTGATCTCTCATCACTTTGAGGCGCACTATCTTTGTTTATAGAAAT 5779  
Db 6755 AAATCAGCTCGTGTGATCTCTCATCACTTTGAGGCGCACTATCTTTGTTTATAGAAAT 6814  
Qy 5780 TTGGGAGATCGATATCGAGAAAAGGTAAGCTGATTTTAAACGTAAGAAATTTATCTCAA 5839  
Db 6815 TTGGGAGATCGATATCGAGAAAAGGTAAGCTGATTTTAAACGTAAGAAATTTATCTCAA 6874  
Qy 5840 GATCTCTGCTCGCGCTTTTGGTATGACCGTGAAACCTCTGACATACGACTCCCG 5899  
Db 6895 GATCTCTGCTCGCGCTTTTGGTATGACCGTGAAACCTCTGACATACGACTCCCG 6934  
Qy 5900 GAGACGCTCACAGCTTGTCTGTAGCGGATCGCGGAGCAGACAAGCCGTCAGGGCGCG 5959  
Db 6935 GAGACGCTCACAGCTTGTCTGTAGCGGATCGCGGAGCAGACAAGCCGTCAGGGCGCG 6994  
Qy 5960 TCAGCGGCTGTGGCGGTGTGGGGCGCGGCATGACCCAGTCACGTAGCGATAGCGGA 6019  
Db 6995 TCAGCGGCTGTGGCGGTGTGGGGCGCGGCATGACCCAGTCACGTAGCGATAGCGGA 7054  
Qy 6020 GTGTATCTGGCTTAACATATGCGGCATCAGAGCAGATTTGTACTGAGAGTGACCATATGC 6079  
Db 7055 GTGTATCTGGCTTAACATATGCGGCATCAGAGCAGATTTGTACTGAGAGTGACCATATGC 7114  
Qy 6080 GGTGTGAATACCGCACAGATCGTAAAGGAAATACCGATCAGGCGCTCTTCCGGTT 6139  
Db 7115 GGTGTGAATACCGCACAGATCGTAAAGGAAATACCGATCAGGCGCTCTTCCGGTT 7174  
Qy 6140 CCGCTCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGAGCGGTATCAGCTCACT 6199  
Db 7175 CCGCTCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGAGCGGTATCAGCTCACT 7234  
Qy 6200 CAAAGCGGTATACGGTTATCCAGAAATCAGGGGTAACGCAAGAAAGCAATGTGAG 6259  
Db 7235 CAAAGCGGTATACGGTTATCCAGAAATCAGGGGTAACGCAAGAAAGCAATGTGAG 7294  
Qy 6260 CAAAGCGGTATACGGTTATCCAGAAATCAGGGGTAACGCAAGAAAGCAATGTGAG 6319  
Db 7295 CAAAGCGGTATACGGTTATCCAGAAATCAGGGGTAACGCAAGAAAGCAATGTGAG 7354  
Qy 6320 GCTCGCGCTCTGACGAGCATCAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 6379  
Db 7355 GCTCGCGCTCTGACGAGCATCAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 7414  
Qy 6380 CGACAGGACTATAAGATACAGGGGTTTCCCGCTGGAAGCTCCCTCGTGCGCTCTCCTG 6439  
Db 7415 CGACAGGACTATAAGATACAGGGGTTTCCCGCTGGAAGCTCCCTCGTGCGCTCTCCTG 7474  
Qy 6440 TTCCGACCTGCGCTTACGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCGTGGCGC 6499  
Db 7475 TTCCGACCTGCGCTTACGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCGTGGCGC 7534  
Qy 6500 TTCTCATAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTCCAGAGTGG 6559  
Db 7535 TTCTCATAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTCCAGAGTGG 7594  
Qy 6560 GCTGTGTACGAAACCCCGTTACGCGGACCGCTGCGCTTATCCCGTAACTATCGTC 6619  
Db 7595 GCTGTGTACGAAACCCCGTTACGCGGACCGCTGCGCTTATCCCGTAACTATCGTC 7654  
Qy 6620 TTGAGTCCAAACCCCGTTAAGACAGCACTTATCGCCACTGGCAGCAGCACTGGTAAACAGGA 6679  
Db 7655 TTGAGTCCAAACCCCGTTAAGACAGCACTTATCGCCACTGGCAGCAGCACTGGTAAACAGGA 7714

Qy 6680 TTAGCAGACGAGGATATGTAGCGGTGTCTACAGAGTTCTTTGAAGTGTGGCCTAACTACG 6739  
Db 7715 TTAGCAGACGAGGATATGTAGCGGTGTCTACAGAGTTCTTTGAAGTGTGGCCTAACTACG 7774  
Qy 6740 GCTACACTAAGAGGACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACTTTGGAA 6799  
Db 7775 GCTACACTAAGAGGACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACTTTGGAA 7834  
Qy 6800 AAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACACCGCTGTAGCGGTGTTTTTGG 6859  
Db 7835 AAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACACCGCTGTAGCGGTGTTTTTGG 7894  
Qy 6860 TTGTGAAGCAGCAGATTTACGCGCAAAAAAAGGATCTCAAGAAGATCTTTTGAATCTTTT 6919  
Db 7895 TTGTGAAGCAGCAGATTTACGCGCAAAAAAAGGATCTCAAGAAGATCTTTTGAATCTTTT 7954  
Qy 6920 CTACGGGCTGTGACGCTCAGTGTGAAGCAAACTCAGCTTTAAGGATTTTGGTCAAGAT 6979  
Db 7955 CTACGGGCTGTGACGCTCAGTGTGAAGCAAACTCAGCTTTAAGGATTTTGGTCAAGAT 8014  
Qy 6980 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAATCAATCT 7039  
Db 8015 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAATCAATCT 8074  
Qy 7040 AAAGTATATATGATGATTAACCTTGTCTGACAGTTTACCAATGCTTAATCAGTAGGCACTA 7099  
Db 8075 AAAGTATATATGATGATTAACCTTGTCTGACAGTTTACCAATGCTTAATCAGTAGGCACTA 8134  
Qy 7100 TCTCAGGATCTCTATTTTGGTTCATCATAGTTGCTTCCCGCTCGTGTAGATAA 7159  
Db 8135 TCTCAGGATCTCTATTTTGGTTCATCATAGTTGCTTCCCGCTCGTGTAGATAA 8194  
Qy 7160 CTACGATACGGGAGGCTTACCCTTGGGCCAGTCTGCAATGATACCGCAGAGCCAC 7219  
Db 8195 CTACGATACGGGAGGCTTACCCTTGGGCCAGTCTGCAATGATACCGCAGAGCCAC 8254  
Qy 7220 GCTCAGCGCTCCAGATTTATCAGCAATAAACCAGCAGCGGAGGCGGAGCGAGAA 7279  
Db 8255 GCTCAGCGCTCCAGATTTATCAGCAATAAACCAGCAGCGGAGGCGGAGCGAGAA 8314  
Qy 7280 GTGTCCTGCAACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTAGAG 7339  
Db 8315 GTGTCCTGCAACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTAGAG 8374  
Qy 7340 TAAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTGCTTGTGAGGATCGTGG 7399  
Db 8375 TAAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTGCTGAGGATCGTGG 8434  
Qy 7400 TGTACGCTCGTGTGTTGGTATGCTTCAATTCAGCTCCGCTTCCCAAGATCAAGCGAG 7459  
Db 8435 TGTACGCTCGTGTGTTGGTATGCTTCAATTCAGCTCCGCTTCCCAAGATCAAGCGAG 8494  
Qy 7460 TTACATCATCCCATCTGTCGCAAAAAAGCGTTAGCTCCTTCCGCTCCTCCGATCGTTG 7519  
Db 8495 TTACATCATCCCATCTGTCGCAAAAAAGCGTTAGCTCCTTCCGCTCCTCCGATCGTTG 8554  
Qy 7520 TCAGAAATAGTTGGCGCAGTGTATCATCTCATGTTATGGCAGCACTGCAATTTCTC 7579  
Db 8555 TCAGAAATAGTTGGCGCAGTGTATCATCTCATGTTATGGCAGCACTGCAATTTCTC 8614  
Qy 7580 TTACTGTTCATGCCATCCGTAAGATGCTTTCTGTGACTGCTGAGTACTCAACCAAGTCAT 7639  
Db 8615 TTACTGTTCATGCCATCCGTAAGATGCTTTCTGTGACTGCTGAGTACTCAACCAAGTCAT 8674  
Qy 7640 TCTGAGAAATAGTGTATCCGCGACCGAGTTGCTCTTCCCGCGGTCAACGCGGATAATA 7699  
Db 8675 TCTGAGAAATAGTGTATCCGCGACCGAGTTGCTCTTCCCGCGGTCAACGCGGATAATA 8734  
Qy 7700 CCGCGCACATAGCAGAACTTTAAAGTGTCTCATCTTGAAGAAAGCTTTCTCGGGCGAA 7759  
Db 8735 CCGCGCACATAGCAGAACTTTAAAGTGTCTCATCTTGAAGAAAGCTTTCTCGGGCGAA 8794

```
QY 7760 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGTGCACCCA 7819
DB 7795 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGTGCACCCA 8854
QY 7820 ACTGATCTTCAGCATCTTTTACTTTCACCGCTTTCTGGGTGAGCAAAAACAGGAAGGC 7879
DB 8855 ACTGATCTTCAGCATCTTTTACTTTCACCGCTTTCTGGGTGAGCAAAAACAGGAAGGC 8914
QY 7880 AAAATGCCGCAAAAAGGGAATAGGCGGACACGGAATGTTGAATACTCATCTCTTCC 7939
DB 8915 AAAATGCCGCAAAAAGGGAATAGGCGGACACGGAATGTTGAATACTCATCTCTTCC 8974
QY 7940 TTTTTCATATATTGAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTG 7999
DB 8975 TTTTTCATATATTGAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTG 9034
QY 8000 AATGTAATTTAGAAAATTAACAAATAGGGGTTCCGGGACATTTCCCGGAAAGTGCCAC 8059
DB 9035 AATGTAATTTAGAAAATTAACAAATAGGGGTTCCGGGACATTTCCCGGAAAGTGCCAC 9094
QY 8060 CTGAGCTCTAAGAAACCAATTTATTCATGACATTAACCTATAAAATAGGCGTATCACGA 8119
DB 9095 CTGAGCTCTAAGAAACCAATTTATTCATGACATTAACCTATAAAATAGGCGTATCACGA 9154
QY 8120 GGCCCTTCGCTCTCAAGAAATTAATTCATGTTTGAAGTTCATGATGATGATGATGATG 8179
DB 9155 GGCCCTTCGCTCTCAAGAAATTAATTCATGTTTGAAGTTCATGATGATGATGATGATG 9214
QY 8180 TCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8239
DB 9215 TCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 9274
QY 8240 CG 8241
DB 9275 CG 9276

RESULT 12
AAD47187
ID AAD47187 standard; DNA; 9276 BP.
AC AAD47187;
XX
DT 24-FEB-2003 (first entry)
XX
DE pPIC9K vector DNA containing human CVSP14 sequence.
XX
KW Serine protease 14; CVSP14; cancer; malignancy; breast; colon;
KW gene therapy; cytostatic; gene; ds.
XX
OS Homo sapiens.
OS Pichia pastoris.
OS Unidentified.
OS Chimeric.
XX
Key Location/Qualifiers
FH 1. 948
FT promoter
FT /tag= a
FT /note= "AOX1 promoter fragment"
FT 855. .875
FT /tag= b
FT /bound_moiety= "AOX1 primer"
FT 949. .1218
FT misc_feature
FT /tag= c
FT /note= "Alpha-factor secretory signal sequence"
FT 1152. .1172
FT /tag= d
FT /bound_moiety= "Alpha-factor primer"
FT 1192. .1241
FT misc_feature
FT /tag= e
FT /note= "Multiple cloning site"
FT 1253. .1586
FT misc_signal
FT /tag= f
```

```
FT primer_bind
FT 1327. .1347
FT /tag= g
FT /bound_moiety= "AOX1 primer"
FT complement(1980. .4514)
FT /tag= h
FT /note= "HIS4 ORF"
FT complement(4928. .5743)
FT misc_feature
FT /tag= i
FT /note= "Kanamycin resistance gene"
FT 6122. .6879
FT /tag= j
FT /note= "AOX1 fragment"
FT complement(7288. .7961)
FT /tag= k
FT /note= "ColE1 origin"
FT complement(8106. .8966)
FT /tag= l
FT /note= "Ampicillin resistance gene"
XX
XX WO200277263-A2.
XX
XX 03-OCT-2002.
XX
XX 20-MAR-2002; 2002WO-US009039.
XX
XX 22-MAR-2001; 2001US-0278166P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Madison EL, Yeh J;
XX
XX WPI; 2003-018938/01.
XX
XX New purified CVSP14 polypeptide and encoding nucleic acid molecule,
XX useful for diagnosing, preventing and/or treating disorders, such as
XX cancers and malignancies of the breast, cervix, prostate, lung, ovary or
XX colon.
XX
XX Example 2; Page 179-191; 185pp; English.
XX
XX The invention relates to transmembrane serine protease 14 (CVSP14), its
XX nucleic acid sequence and the method based on them. The methods and
XX compositions of the invention are useful for diagnosing, preventing
XX and/or treating conditions associated with the aberrant expression or
XX activity of the CVSP14 polypeptide, such as cancers and malignancies of
XX the breast, cervix, prostate, lung, ovary or colon. The methods are also
XX useful for identifying compounds that will modulate the protease activity
XX of CVSP14 polypeptide, and monitoring tumour progression and/or
XX therapeutic effectiveness. CVSP14 DNA used in gene therapy. The present
XX sequence is Pichia pastoris pPIC9K vector DNA containing human CVSP14
XX sequence. This sequence is used in the exemplification of the invention
XX
XX Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;
SQ
Query Match 79.6%; Score 6561.6; DB 7; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;
QY 1 AGATCTAATCAATCCAAAGACGAAAGTTGAATGAACCTTTTGGCATCCGATCCACAG 60
DB 1 AGATCTAATCAATCCAAAGACGAAAGTTGAATGAACCTTTTGGCATCCGATCCACAG 60
QY 61 GTCCATTCTCACATAGTGCCTCAACGCAACGAGGGGATACACTAGACGAGACCGT 120
DB 61 GTCCATTCTCACATAGTGCCTCAACGCAACGAGGGGATACACTAGACGAGACCGT 120
QY 121 TGCATAACGAGGAGCTCCACTCTCTTCTCTCAACACCCACTTTTGGCATTCGAAAACC 180
DB 121 TGCATAACGAGGAGCTCCACTCTCTTCTCTCAACACCCACTTTTGGCATTCGAAAACC 180
QY 181 AGCCCGAGTTATTGGCTTGTGTTGAGCTCGCTCAATTCATTCCTTATTAGCTACTA 240
```





QY 2401 AGGAGTAATCTCCAACTCTCTGGAGATAGGACCAACAAACACAGATCCAGCGTGT 2460  
DB 2195 AGGAGTAATCTCCAACTCTCTGGAGATAGGACCAACAAACACAGATCCAGCGTGT 2254  
QY 2461 GTACTGATCAATCAAGAAAGCAATCTCGATTTGCAGATCAAGTGTTCAGGAGCGT 2520  
DB 2255 GTACTGATCAATCAAGAAAGCAATCTCGATTTGCAGATCAAGTGTTCAGGAGCGT 2314  
QY 2521 ACTGATTTGGACATTTCCAAAGCTGCTCTGATGTTGCAACCGATAGGGTGTAGAGTGT 2580  
DB 2315 ACTGATTTGGACATTTCCAAAGCTGCTCTGATGTTGCAACCGATAGGGTGTAGAGTGT 2374  
QY 2581 CAATACACTCTGCTACAAATTTCAACCTTGGCAACTGCACAGCTTGTGTTGTAACAGCAT 2640  
DB 2375 CAATACACTCTGCTACAAATTTCAACCTTGGCAACTGCACAGCTTGTGTTGTAACAGCAT 2434  
QY 2641 CTTCAATTTCTGGAGCTCTCTGTCTGTCATATCGACAGCCAAACAGAAATCACTTGGGAAT 2700  
DB 2435 CTTCAATTTCTGGAGCTCTCTGTCTGTCATATCGACAGCCAAACAGAAATCACTTGGGAAT 2494  
QY 2701 CAATACCAATTTGAGCTTTGACACAGAGCTCTGAGGCAACGAAATCTGATCAGCGTAT 2760  
DB 2495 CAATACCAATTTGAGCTTTGACACAGAGCTCTGAGGCAACGAAATCTGATCAGCGTAT 2554  
QY 2761 TATCAGCAATAACTAGAACTTTGAGAAGCCCGCAGCAGGCAATGTCAATACTACACAGGCGTG 2820  
DB 2555 TATCAGCAATAACTAGAACTTTGAGAAGCCCGCAGCAGGCAATGTCAATACTACACAGGCGTG 2614  
QY 2821 ATGTGTCATTTGAAACCAATCTTGGCAGCAGTAAAGCACTGTTCTTGGACCAATA 2880  
DB 2615 ATGTGTCATTTGAAACCAATCTTGGCAGCAGTAAAGCACTGTTCTTGGACCAATA 2674  
QY 2881 TTTTGTCACTTAGGACAGTTCTGTTCGGTAAGCCATAGCAGCTACTGCTTGGCGCG 2940  
DB 2675 TTTTGTCACTTAGGACAGTTCTGTTCGGTAAGCCATAGCAGCTACTGCTTGGCGCG 2734  
QY 2941 CTCCTGTAGCAGATACATTTAGCACCAACTTGTGGGCAACGATAGACTTCTTGGGG 3000  
DB 2735 CTCCTGTAGCAGATACATTTAGCACCAACTTGTGGGCAACGATAGACTTCTTGGGG 2794  
QY 3001 TAAAGGTACCACTCTTCTTGGTGGAGATGCAAAACAAATTTCTTTCGACACAGCACTT 3060  
DB 2795 TAAAGGTACCACTCTTCTTGGTGGAGATGCAAAACAAATTTCTTTCGACACAGCACTT 2854  
QY 3061 TGGCAGGAACCCAGCATCAGGAAAGTGGAAAGCAGAAATTCGGGTTCCACAGGAATAT 3120  
DB 2855 TGGCAGGAACCCAGCATCAGGAAAGTGGAAAGCAGAAATTCGGGTTCCACAGGAATAT 2914  
QY 3121 AGAGCCCAACTTCTCAATAGTCTTCGAAACAGAGCAGCTACACAGGCGCAAGTCT 3180  
DB 2915 AGAGCCCAACTTCTCAATAGTCTTCGAAACAGAGCAGCTACACAGGCGCAAGTCT 2974  
QY 3181 CAACTTGCACAGCTCTCCGTTAGTTGAGCTTCATGGAATTTCTGACGTTATCTATAGAGA 3240  
DB 2975 CAACTTGCACAGCTCTCCGTTAGTTGAGCTTCATGGAATTTCTGACGTTATCTATAGAGA 3034  
QY 3241 GATCAATGGCTCTTAAAGTATCTGGCAATTCGATAGTCTCTCTGGGAAAGGAGCTT 3300  
DB 3035 GATCAATGGCTCTTAAAGTATCTGGCAATTCGATAGTCTCTCTGGGAAAGGAGCTT 3094  
QY 3301 CTAACACAGGTGTCTTCAAGCGACTCTCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT 3360  
DB 3095 CTAACACAGGTGTCTTCAAGCGACTCTCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT 3154  
QY 3361 TGTCAACATTTTGACGAAATGTTGCAAAATGTTGATTAATTCATATCTGTTCCG 3420  
DB 3155 TGTCAACATTTTGACGAAATGTTGCAAAATGTTGATTAATTCATATCTGTTCCG 3214  
QY 3421 TTTTCTGATAGGACGAGGAGGCACTCTTCAATTTCTTGTGAGGAGGCGCTTAGAAACGT 3480  
DB 3215 TTTTCTGATAGGACGAGGAGGCACTCTTCAATTTCTTGTGAGGAGGCGCTTAGAAACGT 3274  
QY 3481 CAATTTTGCAAAATTCAAATCGAOCCTTCAGAAAGGAGCTTCTTTAGGTTTGGATTTCTT 3540

DB 3275 CAATTTTGCAAAATTCANATAGACCTTCAGAAAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3334  
QY 3541 TAGGTTGTTCTTGGTGTATCTCTGGTGGGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3600  
DB 3335 TAGGTTGTTCTTGGTGTATCTCTGGTGGGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3394  
QY 3601 ACTTCATATCCAGGTTCTCTCCACTCTGTCACAGCTCACACCGTACTTGGCAGCATCTAA 3660  
DB 3395 ACTTCATATCCAGGTTCTCTCCACTCTGTCACAGCTCACACCGTACTTGGCAGCATCTAA 3454  
QY 3661 CTAATGCAAAATAAAATAAAGTCAGACATTCACAGGCTATATCTTCTTCTTCTTCTTCTTCTT 3720  
DB 3455 CTAATGCAAAATAAAATAAAGTCAGACATTCACAGGCTATATCTTCTTCTTCTTCTTCTTCTT 3514  
QY 3721 CTGCAAGTTTCATCAGCTTCTCTCCCTTAATTTTAGGGTTCAACAAAACCTTCGTCGTCAAATA 3780  
DB 3515 CTGCAAGTTTCATCAGCTTCTCTCCCTTAATTTTAGGGTTCAACAAAACCTTCGTCGTCAAATA 3574  
QY 3781 ACCGTTTGGTATTAAGAACCTTCTGAGCATTTCTTACGATCCCAAGGTCGCTTCCA 3840  
DB 3575 ACCGTTTGGTATTAAGAACCTTCTGAGCATTTCTTACGATCCCAAGGTCGCTTCCA 3634  
QY 3841 TGGCTCTTAAGACCTTTGATTTGGCCAAAACAGGAAGTCGCTTCCAGTGACAGAAACCAA 3900  
DB 3635 TGGCTCTTAAGACCTTTGATTTGGCCAAAACAGGAAGTCGCTTCCAGTGACAGAAACCAA 3694  
QY 3901 CACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACCCA 3960  
DB 3695 CACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACCCA 3754  
QY 3961 GCAACTTTGATGTTCTCCAGTATGACCTTTATACCAAAACCGTGACGAGCAAT 4020  
DB 3755 GCAACTTTGATGTTCTCCAGTATGACCTTTATACCAAAACCGTGACGAGCAAT 3814  
QY 4021 GGTGACCTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGACAGGTACACCA 4080  
DB 3815 GGTGACCTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGACAGGTACACCA 3874  
QY 4081 GGGCCCAACGAGTAATTAAGAGTCAGCCACCAAGTAGTCAATAGACCAATAGACCAAT 4140  
DB 3875 GGGCCCAACGAGTAATTAAGAGTCAGCCACCAAGTAGTCAATAGACCAATAGACCAAT 3934  
QY 4141 CAGTATCAAGACCCCAACAAATTTTCACTGACAGGAACTTTTGGACATCTTCAGAA 4200  
DB 3935 CAGTATCAAGACCCCAACAAATTTTCACTGACAGGAACTTTTGGACATCTTCAGAA 3994  
QY 4201 GTTCTGATTTCACTAGTCAATTCGAGCATCAATAATGGGATTTATACCAAGCAACAG 4260  
DB 3995 GTTCTGATTTCACTAGTCAATTCGAGCATCAATAATGGGATTTATACCAAGCAACAG 4054  
QY 4261 TGGAGTCAATCTACCAACTTTGGGTCAGAAAAGAGCAATACAGTCTTCTACTACGC 4320  
DB 4055 TGGAGTCAATCTACCAACTTTGGGTCAGAAAAGAGCAATACAGTCTTCTACTACGC 4114  
QY 4321 CATTAGTGAACCTTTTCAATTCGCGCAGTCAAGAAAAGGCAACAGCATCTAGCAT 4380  
DB 4115 CATTAGTGAACCTTTTCAATTCGCGCAGTCAAGAAAAGGCAACAGCATCTAGCAT 4174  
QY 4381 TAGCGGCAAGGATCAACTTTTATCAACAGGTCCTATAGATAACCTTAGCGCTGGGA 4440  
DB 4175 TAGCGGCAAGGATCAACTTTTATCAACAGGTCCTATAGATAACCTTAGCGCTGGGA 4234  
QY 4441 TCATCTCTTTGACCACTCTTCTGTCGCAAACTTAGTTCGCAAACTCACTTCTTCTTCTTCTTCTT 4500  
DB 4235 TCATCTCTTTGACCACTCTTCTGTCGCAAACTTAGTTCGCAAACTCACTTCTTCTTCTTCTTCTT 4294  
QY 4501 TATTGTCAACCTTGACAGGTTGTCAGTCACTCTTCAAAATTTGCTCTCTCTTAACGATG 4560  
DB 4295 TATTGTCAACCTTGACAGGTTGTCAGTCACTCTTCAAAATTTGCTCTCTCTTAACGATG 4354  
QY 4561 ACTCAACTTGACATTTAATTTGAGTCAGTCGATTTGATGAACTTTGATCAGGTTGTGCA 4620

4355 ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTGAGTGAACCTTGATCAGGTTGTGCA 4414 Db  
4621 GCTGTGACGACATAGGAAACACCGCTTTTCTACCAAACTCAAGAAATATCAAACT 4680 Qy  
4415 GCTGTGACGACATAGGAAACACCGCTTTTCTACCAAACTCAAGAAATATCAAACT 4474 Db  
4681 CTGCAACACTTTCGCTATGACGATAGCAAGGAAATGTCTATCTTGAAGTCGGAAGTGAAG 4740 Qy  
4475 CTGCAACACTTTCGCTATGACGATAGCAAGGAAATGTCTATCTTGAAGTCGGAAGTGAAG 4534 Db  
4741 TGTAGCTTTGAGAAATTCCTGAAGCCGTATTTTATATATCATGTAGTCACTCATCAGGAGA 4800 Qy  
4535 TGTAGCTTTGAGAAATTCCTGAAGCCGTATTTTATATATCATGTAGTCACTCATCAGGAGA 4594 Db  
4801 TCCTCTACGCGGACGCTCGTGGCGCA- 4828 Qy  
4595 TCCTCTACGCGGACGCTCGTGGCGCACTGACGCGGGGGGGGGGGCGCTGAGGCTGTC 4654 Db  
4829 - 4828 Qy  
4655 CTGCTGAAGAAGGTGTGTGACTCATACAGGCCCTGAATCGCCCATCATCCAGCCAGA 4714 Db  
4829 - 4828 Qy  
4715 AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTGTAGGTGGAACGATGTTGATTTTGA 4774 Db  
4829 - 4828 Qy  
4775 ACTTTTGTCTTCCACGGAACGGTCTCGTGTGCGGAAGATGCGTGATCTGATCCTTCA 4834 Db  
4829 - 4828 Qy  
4835 ACTCAGAAAGTTCGATTTATCAAAAGCCGCTCCGTCAGTCAAGTCAGGTAATGCT 4894 Db  
4829 - 4828 Qy  
4895 CTGCCAGTGTACACCAATTAACCAATTCGATTAGAAAACTCATCGAGCATCAATG 4954 Db  
4829 - 4828 Qy  
4955 AAACTGCAATTTATTCATATCAGGATTTATCAATACCATATTTTGAAAAAGCCGTTTCTG 5014 Db  
4829 - 4828 Qy  
5015 TAATGAAGGAGAAACTCACCGAGGAGTTCATAGGATGGCAAGATCCTGGTATCGTGC 5074 Db  
4829 - 4828 Qy  
5075 TCGATTCCGACTCGTCCAAATCAATCAACCTATTAATTTCCCTCGTCAAAAATAAG 5134 Db  
4829 - 4828 Qy  
5135 GTTATCAAGTGAGAAATCACCATGAGTACGACTGAATCCGGTGAGAAATGGCAAAAGCTT 5194 Db  
4829 - 4828 Qy  
5195 ATGCATTTCTTCCAGACTTGTTCACAGCCGACCATTAACCTCGTCAATCAAAATCACT 5254 Db  
4829 - 4828 Qy  
5255 CGCATCAACCAACCGTTATTCATTGTTGATTCGCTGAGCGAGCGAAATACGCGATC 5314 Db  
4829 - 4828 Qy  
5315 GCTGTTAAAGGACAAATTAACAAACAGGAATCGAATGCAACCGCGCAGGAACACTGCCAG 5374 Db  
4829 - 4828 Qy  
5375 CGCATCAACAATTTTACCTGAATCAGGATATCTTCTAATACCTGGAATGCTGTTT 5434 Db  
4829 - 4828 Qy  
5435 CCCGGGATCGCAGTGGTGAAGTAAACATGCATCATCAGGATCAGGATTAATAATGTTGAT 5494 Db

4829 - 4828 Qy  
5495 GGTGGAAGGAGCATAAATCCGTAGCCAGTTTAGTCTGACCATTCTCATCTCTAACAATC 5554 Db  
4829 - 4828 Qy  
5555 ATTGGCAACGCTACCTTTTGCCATGTTTTCAGAAACAACTCTGGCGCATCGGCTTCCCAT 5614 Db  
4829 - 4828 Qy  
5615 CAATCATAGATTGTGCACTGATTTGCCGACATTTATCGGAGCCCATTTATACCCATA 5674 Db  
4829 - 4828 Qy  
5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGCCCTCGAGAACGCTTTCCGTTGAAT 5734 Db  
4829 - 4828 Qy  
5735 ATGGCTATAAACCCCTTGTATTACTGTTATGTAAGCAGACAGTTTATTGTTTCATGA 5794 Db  
4829 - 4828 Qy  
5795 TGATATATTTTATCTTGTCCATGTAACATCAGAGATTTTGAGACACAACTGGCTTTC 5854 Db  
4829 - 4828 Qy  
5855 CCTGCAAGTGGCATACCGCGGCCACAGGTGCGGTTGCTGGCGCCTATAT 4879 Db  
5855 CCCCCCCCCCTGAGTGGCATACCGCGGCCACAGGTGCGGTTGCTGGCGCCTATAT 5914 Db  
4880 CGCGCATCACCATGAGGGAAGATCGGCTCCGCTCCGCTCATGAGCGCTTGT 4939 Qy  
5915 CGCGCATCACCATGAGGGAAGATCGGCTCCGCTCCGCTCATGAGCGCTTGT 5974 Db  
4940 CGCGCTGGTATGTTGCGAGCCCGCGGCGGAGTGTGGGCGCATCTCCCTGCA 4999 Qy  
5975 CGCGTGGTATGTTGCGAGCCCGCGGCGGAGTGTGGGCGCATCTCCCTGCA 6034 Db  
5000 TGACACCATCTCTTCCGCGCGGCTGCTCAACGCGCTCACTACTACTGCGCTTCTCT 5059 Qy  
6035 TGACACCATCTCTTCCGCGCGGCTGCTCAACGCGCTCACTACTACTGCGCTTCTCT 6094 Db  
5060 AATGCAAGGATGTCATGAAGGAGAGGCTCGAGTATCTATGATTGGAAGTATGGGAATGGT 5119 Qy  
6095 AATGCAAGGATGTCATGAAGGAGAGGCTCGAGTATCTATGATTGGAAGTATGGGAATGGT 6154 Db  
5120 GATACCGCATCTCTGATGCTGTTGAGTCTCTATCAGATTTATGCCCACTAAGCAAC 5179 Qy  
6155 GATACCGCATCTCTGATGCTGTTGAGTCTCTATCAGATTTATGCCCACTAAGCAAC 6214 Db  
5180 CGGAGGAGGAGATTTTCATGTTAAATTTCTGACTTTTGTCTCATCAGTAGACTCGAACTG 5239 Qy  
6215 CGGAGGAGGAGATTTTCATGTTAAATTTCTGACTTTTGTCTCATCAGTAGACTCGAACTG 6274 Db  
5240 TGAGACTATCTCGGTTATGACAGCAAAATGTCCTTCTTGGAGACAGTAAATGAAGTCCC 5299 Qy  
6275 TGAGACTATCTCGGTTATGACAGCAAAATGTCCTTCTTGGAGACAGTAAATGAAGTCCC 6334 Db  
5300 ACCAATAAGAAATCCTTGTATCAGAAACAACTTCTTGTTCGAACTTTTTCGTTGCC 5359 Qy  
6335 ACCAATAAGAAATCCTTGTATCAGAAACAACTTCTTGTTCGAACTTTTTCGTTGCC 6394 Db  
5360 TTGAACATATAAATGTAGAGTGGATATGTCGGTAGGAATGGAGCGGCGAAATGCTTACC 5419 Qy  
6395 TTGAACATATAAATGTAGAGTGGATATGTCGGTAGGAATGGAGCGGCGAAATGCTTACC 6454 Db  
5420 TTCTGACCTTCAAGAGTATGAGGTTGTTAGTACTGATGATGCACTTCACTGAGCAAC 5479 Qy  
6455 TTCTGACCTTCAAGAGTATGAGGTTGTTAGTACTGATGATGCACTTCACTGAGCAAC 6514 Db  
5480 GTTGTATTTTCTGTTCAAACTTCCGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATT 5539 Qy  
6515 GTTGTATTTTCTGTTCAAACTTCCGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATT 6574 Db



```

Db 8735 CCGGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAACCGTTCTTCGGGGCGAA 8794
Qy 7760 AACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTCGATGTAAACCCACCTCGTGCCACCA 7819
Db 8795 AACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTCGATGTAAACCCACCTCGTGCCACCA 8854
Qy 7820 ACTGATCTTCAGCATCTTTTACTTTTCCACCGCTTCTCGGTGAGCAAAAACAGGAAGGC 7879
Db 8855 ACTGATCTTCAGCATCTTTTACTTTTCCACCGCTTCTCGGTGAGCAAAAACAGGAAGGC 8914
Qy 7880 AABATGCCGAAAAGGATAGGGCGACACGGAATGTGAATCTCATCTCTTCC 7939
Db 8915 AABATGCCGAAAAGGATAGGGCGACACGGAATGTGAATCTCATCTCTTCC 8974
Qy 7940 TTTTTCATATTTATTGAAGCATTTTATCAGGGTTTCTCTCATGAGCGGATACATATTG 7999
Db 8975 TTTTTCATATTTATTGAAGCATTTTATCAGGGTTTCTCTCATGAGCGGATACATATTG 9034
Qy 8000 AATGATTTAGAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAAAGTGCCAC 8059
Db 9035 AATGATTTAGAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAAAGTGCCAC 9094
Qy 8060 CTGACGCTTAGAACAACATTTATTCATGACATTAACCTATAAAATAGGCGTATCAGCA 8119
Db 9095 CTGACGCTTAGAACAACATTTATTCATGACATTAACCTATAAAATAGGCGTATCAGCA 9154
Qy 8120 GGCCCTTTCGCTTCAAGAAATAATCTCATGTTTGCAGCTTATCATGATAGCTGAC 8179
Db 9155 GGCCCTTTCGCTTCAAGAAATAATCTCATGTTTGCAGCTTATCATGATAGCTGAC 9214
Qy 8180 TCATGTTGGTATTGTGAATAGACGAGATCGGGAACACTGAAATAAACAGTTATTATT 8239
Db 9215 TCATGTTGGTATTGTGAATAGACGAGATCGGGAACACTGAAATAAACAGTTATTATT 9274
Qy 8240 CG 8241
Db 9275 CG 9276

RESULT 13
ID ABZ58513 standard; DNA; 9276 BP.
XX AC ABZ58513;
XX DT 13-MAY-2003 (first entry)
XX DE Pichia pastoris plasmid pPIC9k.
XX KW Transmembrane serine protease 20; MTS20; enzyme; endothelial;
XX KW cytosolic; dermatological; cardiant; vulnary; ophthalmological;
XX KW gene therapy; pPIC9k; ds.
XX OS Pichia pastoris.
XX FH Key
XX FT promoter Location/Qualifiers
XX FT 1..948
XX FT /tag= a
XX FT /note= "AOX1 promoter fragment"
XX FT 949..1218
XX FT /tag= b
XX FT /note= "alpha-factor secretion signal"
XX FT 1253..1586
XX FT /tag= c
XX FT /note= "AOX1 transcription termination region"
XX FT complement(1980..4514)
XX FT /tag= d
XX FT /product= "HIS4"
XX FT complement(4928..5743)
XX FT /tag= e
XX FT /note= "kanamycin resistance gene"
XX FT 6122..6879
XX FT /tag= f
XX FT misc_feature

```

```

FT misc_feature /note= "3' AOX1 fragment"
FT complement(7288..7961)
FT /tag= g
FT /note= "ColE1 origin"
FT 8106..8966
FT /tag= h
FT /note= "ampicillin resistance gene"
XX PN WC2003004681-A2.
XX 16-JAN-2003.
XX PD 03-JUL-2002; 2002WO-US021208.
XX PF 03-JUL-2001; 2001US-0302939P.
XX PR (CORV-) CORVAS INT INC.
XX PA Madison EL, Ong EO;
XX PI WPI; 2003-239207/23.
XX DR
XX PS Example 2; Page 206-208; 216pp; English.
XX CC The present sequence is that of Pichia pastoris vector plasmid pPIC9k.
CC The vector was used in the expression of cDNA (see ABZ58498) encoding the
CC protease domain (see ABP72374) of novel human type II transmembrane
CC serine protease 20 (MTSP20) in P. pastoris host cells. The invention
CC provides MTS20 polypeptides (see also ABP72374), the nucleic acids
CC encoding them (see also ABZ58499), vectors, host cells, transgenic
CC animals, probes, primers and antibodies. Also claimed are antisense
CC oligonucleotides, double-stranded RNAs and antibodies, and their use in
CC claimed methods of inhibiting tumour initiation, growth or progression,
CC and of treating or preventing a disease or disorder associated with
CC undesired and/or uncontrolled angiogenesis or neovascularisation,
CC especially undesired angiogenesis associated with solid neoplasms,
CC vascular malformations and cardiovascular disorders, chronic inflammatory
CC diseases, aberrant wound repairs, circulatory disorders, crest syndromes,
CC dermatological disorders and ocular disorders
XX SQ Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;

Query Match 79.6%; Score 6561.6; DB 7; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

Qy 1 AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCATCCGATCCACAG 60
Db 1 AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCATCCGATCCACAG 60
Qy 61 GTCCATTCTCACATAAGTGCACCAACGACAGGGGGATACACTAGCAGACCGCT 120
Db 61 GTCCATTCTCACATAAGTGCACCAACGACAGGGGGATACACTAGCAGACCGCT 120
Qy 121 TGCACACGAGGAGACCTCCACTCTCTCTCTCAACACACCCACTTTTGGCATCGAAAACC 180
Db 121 TGCACACGAGGAGACCTCCACTCTCTCTCTCAACACACCCACTTTTGGCATCGAAAACC 180
Qy 181 AGCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCTCAATTCCTTCTATTAGGCTACTA 240
Db 181 AGCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCTCAATTCCTTCTATTAGGCTACTA 240
Qy 241 ACACCATGACTTTATTAGCTCTGTCTATCTGCGCCCTCGGAGGTTTCATGTTTCTTTA 300
Db 241 ACACCATGACTTTATTAGCTCTGTCTATCTGCGCCCTCGGAGGTTTCATGTTTCTTTA 300
Qy 301 TTTCGGAATGCAACAAGCTCCGATTCACCCGAAACATCACTCCAGATGAGGCTTTCTG 360
Db 301 TTTCGGAATGCAACAAGCTCCGATTCACCCGAAACATCACTCCAGATGAGGCTTTCTG 360

```

New type-II membrane-type serine protease 20 polypeptides, useful for preparing a medicament for diagnosing, treating or preventing cancer, dermatological disorders, aberrant wound repairs or crest syndromes.

QY	361	AGTGTGGGTC	AAATAGTTT	CATGTTCC	CCAAATG	CCCAAACT	GACAGTTT	AAACGGT	420	QY	1441	CGGCGCGAA	TTAAAT	CGCCTT	AGACATG	ACTGTT	CTCAGTT	CAAGT	TGGGCACTT	ACG	1500
Db	361	AGTGTGGGTC	AAATAGTTT	CATGTTCC	CCAAATG	CCCAAACT	GACAGTTT	AAACGGT	420	Db	1235	CGGCGCGAA	TTAAAT	CGCCTT	AGACATG	ACTGTT	CTCAGTT	CAAGT	TGGGCACTT	ACG	1294
QY	421	GTCTTGGAA	CTAATAT	GACAAAG	CGTGAT	CTCATC	CAAGATG	AACTAAGTT	TGTTTCG	480	QY	1501	AGAAGACCG	GTCTG	CTAGATT	CTTAAT	CAAGAGAT	CTCAGA	ATGCCAT	TTCGCTG	1560
Db	421	GTCTTGGAA	CTAATAT	GACAAAG	CGTGAT	CTCATC	CAAGATG	AACTAAGTT	TGTTTCG	480	Db	1295	AGAAGACCG	GTCTG	CTAGATT	CTTAAT	CAAGAGAT	CTCAGA	ATGCCAT	TTCGCTG	1354
QY	481	TTGAAATG	CTTAAC	CGGCGAG	TGTTG	TCAAA	AGAACT	TCBAA	AGTCCG	CTATCCGTTTGT	540	QY	1561	ATGACAGG	CTTCAT	TTTTTG	ATACCT	TATATG	TAAGTAT	AGGAT	1620
Db	481	TTGAAATG	CTTAAC	CGGCGAG	TGTTG	TCAAA	AGAACT	TCBAA	AGTCCG	CTATCCGTTTGT	540	Db	1355	ATGACAGG	CTTCAT	TTTTTG	ATACCT	TATATG	TAAGTAT	AGGAT	1414
QY	541	CTTGTGTT	GGTATG	ATGAG	GAATG	CTCAAAA	ATAAT	TCTCAT	TAAATG	CTTAGCGAGTCT	600	QY	1621	GTCAATTT	GTCTT	CTGAC	AGCTT	GTCTT	CGATC	CGCTAT	1680
Db	541	CTTGTGTT	GGTATG	ATGAG	GAATG	CTCAAAA	ATAAT	TCTCAT	TAAATG	CTTAGCGAGTCT	600	Db	1415	GTCAATTT	GTCTT	CTGAC	AGCTT	GTCTT	CGATC	CGCTAT	1474
QY	601	CTCTAT	CGCTT	CGAA	CCCGG	TGCAC	CTGTG	CGGAA	ACGCAAT	GGGNAACACCGCT	660	QY	1681	TATCTT	GTGGT	AGGGT	TTGGG	AAAAAT	CAAT	TCGAGT	1740
Db	601	CTCTAT	CGCTT	CGAA	CCCGG	TGCAC	CTGTG	CGGAA	ACGCAAT	GGGNAACACCGCT	660	Db	1475	TATCTT	GTGGT	AGGGT	TTGGG	AAAAAT	CAAT	TCGAGT	1534
QY	661	TTTTGG	ATGATT	ATG	CTT	CTC	CAAT	TGTAT	GTCTT	CCAAAGAT	720	QY	1741	ACTCTC	TTTCAG	ATACAG	AAGTAA	AGTGA	AGTTT	CGTTT	1800
Db	661	TTTTGG	ATGATT	ATG	CTT	CTC	CAAT	TGTAT	GTCTT	CCAAAGAT	720	Db	1535	ACTCTC	TTTCAG	ATACAG	AAGTAA	AGTGA	AGTTT	CGTTT	1594
QY	721	GCTGAT	AGCCT	TAAG	CTT	CATG	ATCAAA	ATTTAA	CTT	TAAGCTT	780	QY	1801	GCTTTA	ATGCG	TAGTTT	TATC	CAGTTT	AAAT	TTGCTA	1860
Db	721	GCTGAT	AGCCT	TAAG	CTT	CATG	ATCAAA	ATTTAA	CTT	TAAGCTT	780	Db	1595	GCTTTA	ATGCG	TAGTTT	TATC	CAGTTT	AAAT	TTGCTA	1654
QY	781	ATAA	TAACAG	AGAA	AGT	CGCT	GTCTT	AAAC	CTTTT	TATCAT	840	QY	1861	ATCTAA	CAAT	CGCT	TCAT	CGGCA	CCGT	CA	1920
Db	781	ATAA	TAACAG	AGAA	AGT	CGCT	GTCTT	AAAC	CTTTT	TATCAT	840	Db	1655	ATCTAA	CAAT	CGCT	TCAT	CGGCA	CCGT	CA	1714
QY	841	ACTTTCA	TAAAT	TGG	ACT	GTGTT	CCA	ATGTA	CAAG	CTTTT	900	QY	1921	CTTGT	ATGCG	GTACT	CGCG	CCCTT	CGCG	GAAT	1980
Db	841	ACTTTCA	TAAAT	TGG	ACT	GTGTT	CCA	ATGTA	CAAG	CTTTT	900	Db	1715	CTTGT	ATGCG	GTACT	CGCG	CCCTT	CGCG	GAAT	1774
QY	901	CAACT	TGAG	AGAT	CAAAA	CAACT	TAAT	TTCG	AGAGT	CCAA	960	QY	1981	CAGT	CAC	TATG	CGCT	GTCT	ATG	CAAT	2040
Db	901	CAACT	TGAG	AGAT	CAAAA	CAACT	TAAT	TTCG	AGAGT	CCAA	960	Db	1775	CAGT	CAC	TATG	CGCT	GTCT	ATG	CAAT	1834
QY	961	TCAAT	TTTT	TACT	GCAG	TTT	ATTC	CG	AGCAT	CTCT	1020	QY	2041	TCT	CG	AGCAT	CTCT	CG	CCG	CCG	2100
Db	961	TCAAT	TTTT	TACT	GCAG	TTT	ATTC	CG	AGCAT	CTCT	1020	Db	1835	TCT	CG	AGCAT	CTCT	CG	CCG	CCG	1894
QY	1021	ACA	CAG	AGAT	GAA	ACG	CAAA	TTT	CGG	CTG	1080	QY	2101	AG	CACT	AT	CG	AGAT	CA	CCG	2160
Db	1021	ACA	CAG	AGAT	GAA	ACG	CAAA	TTT	CGG	CTG	1080	Db	1895	AG	CACT	AT	CG	AGAT	CA	CCG	1954
QY	1081	GAAG	GGAT	TT	CGAT	GT	GTG	TTT	GC	CAAT	1140	QY	2161	AAT	G	TAAGT	TAAT	TAAT	TAAT	TAAT	2220
Db	1081	GAAG	GGAT	TT	CGAT	GT	GTG	TTT	GC	CAAT	1140	Db	1955	AAT	G	TAAGT	TAAT	TAAT	TAAT	TAAT	2014
QY	1141	TTT	ATA	ATA	ATA	ATA	ATA	ATA	ATA	ATA	1200	QY	2221	CAG	CA	TTT	CGG	TAG	CA	TTT	2280
Db	1141	TTT	ATA	ATA	ATA	ATA	ATA	ATA	ATA	ATA	1200	Db	2015	CAG	CA	TTT	CGG	TAG	CA	TTT	2074
QY	1201	AG	AG	GGT	GA	AG	GGT	GA	AG	GGT	1260	QY	2281	TAT	GT	TT	CAG	T	CGT	TT	2340
Db	1201	AG	AG	GGT	GA	AG	GGT	GA	AG	GGT	1260	Db	2075	TAT	GT	TT	CAG	T	CGT	TT	2134
QY	1261	CCCC	CCG	GAG	AG	ATAA	TTG	TG	GT	TTCC	1320	QY	2341	TG	TT	AACT	CCG	CTG	AT	CG	2400
Db	1218	-----	-----	-----	-----	-----	-----	-----	-----	-----	1217	Db	2135	TG	TT	AACT	CCG	CTG	AT	CG	2194
QY	1321	AG	AG	TTG	CT	TTT	GT	ATG	AC	AGT	1380	QY	2401	AG	GA	TAAT	CT	CA	CA	CT	2460
Db	1218	-----	-----	-----	-----	-----	-----	-----	-----	-----	1217	Db	2195	AG	GA	TAAT	CT	CA	CA	CT	2254
QY	1381	AT	CG	AG	AA	CA	CT	CA	AG	AG	1440	QY	2461	GT	AC	TT	GT	AT	CA	CA	2520
Db	1218	-----	-----	-----	-----	-----	-----	-----	-----	-----	1217	Db	2255	GT	AC	TT	GT	AT	CA	CA	2314
QY	1440	-----	-----	-----	-----	-----	-----	-----	-----	-----	1234	QY	2521	ACT	GA	TT	GG	CA	TT	TC	2580





QY 4741 TGTAGTCTTGAGAAATCTGAAGCCGATTTTATTATCAGTCAGTCAGTCATCAGGAGA 4800  
Db 4535 TGTAGTCTTGAGAAATCTGAAGCCGATTTTATTATCAGTCAGTCAGTCATCAGGAGA 4594  
QY 4801 TCCTCTACGGCGGAGCATCGTGCCGA----- 4828  
Db 4595 TCCTCTACGGCGGAGCATCGTGCCGACCTGCAGGGGGGGGGGGCGCTGAGGCTGCG 4654  
QY 4829 ----- 4828  
Db 4655 CTCGTGAAGAAGGTGTGCTGACTCATCAGGCGCTGAATCGGCCCATCATCCAGCCAGA 4714  
QY 4829 ----- 4828  
Db 4715 AAGTGAGGGAGCCAGGTGTGATGAGAGCTTTGTGTAGTGGACCAAGTTGGTGAATTTGA 4774  
QY 4829 ----- 4828  
Db 4775 ACTTTGCTTGGCCAGGAACGGTCTGCGTTGTGCGGAAGATCGGTGATCTGATCCCTTCA 4834  
QY 4829 ----- 4828  
Db 4835 ACTCAGCAAAAGTTCGATTTATTCAACAAAGCCGCTCCGTCAGTCAGCGTAATGCT 4894  
QY 4829 ----- 4828  
Db 4895 CTGCCAGTGTTAACAACCAATTAACCAATCTGATTAGAAAAAATCATCGAGCATCAAAATG 4954  
QY 4829 ----- 4828  
Db 4955 AAACTCGAATTTATTCATATCAGGATTAATCAATACCATAATTTTGAAGAGCCGTTTCTG 5014  
QY 4829 ----- 4828  
Db 5015 TAATGAAGGAGAAAACTCACCGAGGCAGTTCATPAGGATGGCAAGATCCTGTGATCGGTC 5074  
QY 4829 ----- 4828  
Db 5075 TGGATTTCCGACTCGTGAACATCAATACACCTATTAAATTCGCCCTGCTCAAAATAG 5134  
QY 4829 ----- 4828  
Db 5135 GTTATCAAGTGAGAAATCACCATCAGTCAGTCAGTGAATCCGGTGAGAAATGGCAAAAGCTT 5194  
QY 4829 ----- 4828  
Db 5195 ATGCATTTCTTCAGACTGTTTCAACAGGCCAGCCATTACGTCGTCATCAAAATCACT 5254  
QY 4829 ----- 4828  
Db 5255 CGCATCAACCAACCGTTATTTCATTGCTGATTTGGCTGAGCGAGACGAAATACGGGATC 5314  
QY 4829 ----- 4828  
Db 5315 GCTGTTAAAGGACAATTAACAACAGGAATGGAATGCAACCGGCGGAGGAACACTGCCAG 5374  
QY 4829 ----- 4828  
Db 5375 CGCATCAACATATTTTACCTGATCAGGATATCTTCTTAATACCTGGAATGCTGTTTT 5434  
QY 4829 ----- 4828  
Db 5435 CCGGGGATCGAGTGTGTAGTAACCATGCATCATCAGGATACGGATAAAATGCTTTGAT 5494  
QY 4829 ----- 4828  
Db 5495 GGTGGAAGAGGCAATAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTACATC 5554  
QY 4829 ----- 4828  
Db 5555 ATTGCAACGCTACTCTTTGGCCATGTTTCAGAAACAACACTCTGGCGCATCGGGCTTCCATA 5614

QY 4829 ----- 4828  
Db 5615 CAATCGATAGATTGTCGACCTGATTGCCGACATTATCGGAGCCCAATTATATACCCATA 5674  
QY 4829 ----- 4828  
Db 5675 TAAATCAGCATCCATGTTGGAATTTAATCGGGCCTCGAGCAAGACGTTTCCCGTTGAAT 5734  
QY 4829 ----- 4828  
Db 5735 ATGGCTCATACACCCCTTGATTACTGTTTATGTAAGCAGACAGTTTATTATTGTTATGA 5794  
QY 4829 ----- 4828  
Db 5795 TGATATAATTTTATCTTGTGCAATGTAATCATCAGAGATTTTGAGACACAAGTGGCTTTC 5854  
QY 4829 ----- 4828  
Db 5855 CCCCCCCCCCTGCAGGTTCGCAATCACCGGCGCCACAGGTGCGGTTGCTGGCGGCTATAT 5914  
QY 4880 CGCGACATCACCGATGGGGAAGATCGGGCTCGCACTTCGGGCTCATGAGCGCTTCTTT 4939  
Db 5915 CGCGACATCACCGATGGGGAAGATCGGGCTCGCACTTCGGGCTCATGAGCGCTTCTTT 5974  
QY 4940 CGGCGTGGGTATGTTGGCAGGCCCGCTGGCGGGGAGTGTGGCGGCCCACTCTCTTGCA 4999  
Db 5975 CGGCGTGGGTATGTTGGCAGGCCCGCTGGCGGGGACTGTTGGCGGCCATCTCTTGCA 6034  
QY 5000 TGCAACATTCCTTGGCGGGCGGTGCTCAACGGCTCAACCTACTACTGCGGCTGCTTCT 5059  
Db 6035 TGCAACATTCCTTGGCGGGCGGTGCTCAACGGCTCAACCTACTACTGCGGCTGCTTCT 6094  
QY 5060 AATCAGGAGTCGATGAAGGAGAGCGTCGAGTATCTATGATCGAAGTATGGGAATGGT 5119  
Db 6095 AATCAGGAGTCGATGAAGGAGAGCGTCGAGTATCTATGATCGAAGTATGGGAATGGT 6154  
QY 5120 GATACCCCAATCTTCACTGTCCTTGAAGTCTCTATCAGATTTATGCCCAACTAAAGCAAC 5179  
Db 6155 GATACCCCAATCTTCACTGTCCTTGAAGTCTCTATCAGATTTATGCCCAACTAAAGCAAC 6214  
QY 5180 CGGAGGAGAGATTCATGTTAAATTTCTGACTTTTGTGTCATCAGTACTGAACTG 5239  
Db 6215 CGGAGGAGAGATTCATGTTAAATTTCTGACTTTTGTGTCATCAGTACTGAACTG 6274  
QY 5240 TGAGACTATCTCGGTTATGACAGAGAAATGTCCTTCTTGGAGACAGTAAATGAAGTCCC 5299  
Db 6275 TGAGACTATCTCGGTTATGACAGAGAAATGTCCTTCTTGGAGACAGTAAATGAAGTCCC 6334  
QY 5300 ACCAATAAGAAATCTTGTATCAGGAAACAACTTCTGTTTTCGAACTTTTCGGTGCC 5359  
Db 6335 ACCAATAAGAAATCTTGTATCAGGAAACAACTTCTGTTTTCGAACTTTTCGGTGCC 6394  
QY 5360 TTGAACATAAAATCTAGAGTGGATATCTCGGTAGGAATGGAGGGGCAAAATGCTTACC 5419  
Db 6395 TTGAACATAAAATCTAGAGTGGATATCTCGGTAGGAATGGAGGGGCAAAATGCTTACC 6454  
QY 5420 TTCTGGAACCTTCAAGAGTATGTAGGTTTGTATGATCTGATGCGCACTTCAGTGACAAAC 5479  
Db 6455 TTCTGGAACCTTCAAGAGTATGTAGGTTTGTATGATCTGATGCGCACTTCAGTGACAAAC 6514  
QY 5480 GTTGTATTTTCGTTCAAAACCAATTCGAATCCAGAGAAATCAAAGTTGTTTCTACTATT 5539  
Db 6515 GTTGTATTTTCGTTCAAAACCAATTCGAATCCAGAGAAATCAAAGTTGTTTCTACTATT 6574  
QY 5540 GATCCAAACCGAGTCGGTCTTTGAAACTGCAATAGTGTGCTCGTGTGTTTGAAGTCACTTT 5599  
Db 6575 GATCCAAACCGAGTCGGTCTTTGAAACTGCAATAGTGTGCTCGTGTGTTTGAAGTCACTTT 6634  
QY 5600 TGTATGAATAATCTAGTCTTTTGAATAAATCTTGAACGAGCAAGGGGCAATAAATACC 5659  
Db 6635 TGTATGAATAATCTAGTCTTTTGAATAAATCTTGAACGAGCAAGGGGCAATAAATACC 6694  
QY 5660 CAAATCTAAAACTCTTTTAAAAACGTTTAAAGGACAAGTATGCTGTGCTGTATTAAAAACCC 5719

Db	6695	CAAACTCTAAACCTCTTTAAACGCTTAAAGGCAAGTATGTCTGCTGTATTAAACCCC	6754
Qy	5720	AAATCAGCTCGTACTGATCCTCATCAACTTGAAGGGGCACTATCTTTGTTTGAAGAAAT	5779
Db	6755	AAATCAGCTCGTACTGATCCTCATCAACTTGAAGGGGCACTATCTTTGTTTGAAGAAAT	6814
Qy	5780	TTGGGGAGATCGGATATCGAGAAAAGGTAGCTGTATTTAAAGCTGAATTTATCTCAA	5839
Db	6815	TTGGGGAGATCGGATATCGAGAAAAGGTAGCTGTATTTAAAGCTGAATTTATCTCAA	6874
Qy	5840	GATCTCTCCCTCGCGGCTTTCCGGTGATGACGGTGAAAACCTCTGACACATGCGACTCCCG	5899
Db	6875	GATCTCTCCCTCGCGGCTTTCCGGTGATGACGGTGAAAACCTCTGACACATGCGACTCCCG	6934
Qy	5900	GAGACGGTCAACAGCTGTCTGTAAAGCGGATGCGGGAGAGACAAAGCCGCTGAGGGGCGG	5959
Db	6935	GAGACGGTCAACAGCTGTCTGTAAAGCGGATGCGGGAGAGACAAAGCCGCTGAGGGGCGG	6994
Qy	5960	TCAGCGGCTGTGGGGGTGTCGGGGCGCAGCATGACCCAGTCACGCTAGCGATAGCGGA	6019
Db	6995	TCAGCGGCTGTGGGGGTGTCGGGGCGCAGCATGACCCAGTCACGCTAGCGATAGCGGA	7054
Qy	6020	GTGTATCTGGCTTAACCTATCGGCATCAGACGAGATTGTACTGAGAGTGACCATATGC	6079
Db	7055	GTGTATCTGGCTTAACCTATCGGCATCAGACGAGATTGTACTGAGAGTGACCATATGC	7114
Qy	6080	GGTGTAAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAGCGCTCTTCCGCTT	6139
Db	7115	GGTGTAAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAGCGCTCTTCCGCTT	7174
Qy	6140	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCACT	6199
Db	7175	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCACT	7234
Qy	6200	CAAGCGGCTAATACGGTTATCCACAGAAATCAGGGGATAACCGGAGAAACATGTGAG	6259
Db	7235	CAAGCGGCTAATACGGTTATCCACAGAAATCAGGGGATAACCGGAGAAACATGTGAG	7294
Qy	6260	CAAAAGGCGCAGAAAGCCAGAAACCGTAAAGGCGCGCTGCTGCGCTTTTCCATA	6319
Db	7295	CAAAAGGCGCAGAAAGCCAGAAACCGTAAAGGCGCGCTGCTGCGCTTTTCCATA	7354
Qy	6320	GGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC	6379
Db	7355	GGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC	7414
Qy	6380	CGACAGGACTATAAAGATACAGGGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG	6439
Db	7415	CGACAGGACTATAAAGATACAGGGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG	7474
Qy	6440	TTCCGACCTGCGGCTTACCGGATACCTGTCCGCTTTTCCCTTCCGGAAGCGTGGCGC	6499
Db	7475	TTCCGACCTGCGGCTTACCGGATACCTGTCCGCTTTTCCCTTCCGGAAGCGTGGCGC	7534
Qy	6500	TTTCTCATAGCTACGCTGTAGGTATCTCAGTTCCGTTAGGTCTGCTCCAAAGCTGG	6559
Db	7535	TTTCTCATAGCTACGCTGTAGGTATCTCAGTTCCGTTAGGTCTGCTCCAAAGCTGG	7594
Qy	6560	GCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCTAACTACGTC	6619
Db	7595	GCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCTAACTACGTC	7654
Qy	6620	TTGAGTCCAAACCCCGTTAAGACAGACTTATGCGCACTGCGACGACCACTGTGTACAGA	6679
Db	7655	TTGAGTCCAAACCCCGTTAAGACAGACTTATGCGCACTGCGACGACCACTGTGTACAGA	7714
Qy	6680	TTAGCAGAGCGGATATGAGGGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACCTAGC	6739
Db	7715	TTAGCAGAGCGGATATGAGGGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACCTAGC	7774
Qy	6740	GCTACACTAAGACAGATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAA	6799

Db	7775	GCTACACTAAGACAGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA	7834
Qy	6800	AAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGCTAGCGGTGTTTTTTTG	6859
Db	7835	AAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGCTAGCGGTGTTTTTTTG	7894
Qy	6860	TTTGCAAGCAGCAGATTTACCGCAGAAAAAAGGATCTCAAGAAGATCTTTGATCTTTT	6919
Db	7895	TTTGCAAGCAGCAGATTTACCGCAGAAAAAAGGATCTCAAGAAGATCTTTGATCTTTT	7954
Qy	6920	CTACGGGGTCTGACGCTGAGTGAAGCAAACTCAGTTTAAGGGATTTTGGTCATGAGAT	6979
Db	7955	CTACGGGGTCTGACGCTGAGTGAAGCAAACTCAGTTTAAGGGATTTTGGTCATGAGAT	8014
Qy	6980	TATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAAAATGAAGTTTAAATCAATCT	7039
Db	8015	TATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAAAATGAAGTTTAAATCAATCT	8074
Qy	7040	AAAGTATATATGAGTAAATTTGGTCTGACAGTTTCAATGCTTAATCATGAGGACCTTA	7099
Db	8075	AAAGTATATATGAGTAAATTTGGTCTGACAGTTTCAATGCTTAATCATGAGGACCTTA	8134
Qy	7100	TCTCAGGATCTGCTTATTTTGGTTCATCCATAGTTGCTGCTGCTGCTGCTGCTGCTGCT	7159
Db	8135	TCTCAGGATCTGCTTATTTTGGTTCATCCATAGTTGCTGCTGCTGCTGCTGCTGCTGCT	8194
Qy	7160	CTACGATACGGGAGGGCTTACCATCTGGCCCCCAGTCTCAATGATACCGCAGACCCAC	7219
Db	8195	CTACGATACGGGAGGGCTTACCATCTGGCCCCCAGTCTCAATGATACCGCAGACCCAC	8254
Qy	7220	GCTACCGGCTCCAGATTTTATCAGCAATAAACCAGCCAGCCGAGGGCCGAGCGAGAA	7279
Db	8255	GCTACCGGCTCCAGATTTTATCAGCAATAAACCAGCCAGCCGAGGGCCGAGCGAGAA	8314
Qy	7280	GTGGTCTCTCAACTTTTATCGGCTCCATCCAGTCTTAAATTTGTTCCCGGAGCTAGAG	7339
Db	8315	GTGGTCTCTCAACTTTTATCGGCTCCATCCAGTCTTAAATTTGTTCCCGGAGCTAGAG	8374
Qy	7340	TAAAGTAGTTCGCGAGTTATAGTTTGGCGAAAGTTGCTGCTGCTGCTGCTGCTGCTGCT	7399
Db	8375	TAAAGTAGTTCGCGAGTTATAGTTTGGCGAAAGTTGCTGCTGCTGCTGCTGCTGCTGCT	8434
Qy	7400	TGTCAGCTCGTCTGTTGGTATGCTTCAATCAGCTCCGCTTCCCAACCATCAAGCGAG	7459
Db	8435	TGTCAGCTCGTCTGTTGGTATGCTTCAATCAGCTCCGCTTCCCAACCATCAAGCGAG	8494
Qy	7460	TTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTTCTGCTGCTTCCGATCGTTG	7519
Db	8495	TTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTTCTGCTGCTTCCGATCGTTG	8554
Qy	7520	TCAGAGTAAAGTTGGCGGAGTGTATCACTCATGCTTATGGCAGCACTGCAATAATCTC	7579
Db	8555	TCAGAGTAAAGTTGGCGGAGTGTATCACTCATGCTTATGGCAGCACTGCAATAATCTC	8614
Qy	7580	TTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCAT	7639
Db	8615	TTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCAT	8674
Qy	7640	TCTGAGAAATAGTGTATCGGGGAGCGAGTTGCTCTTCCCGCGGCTCAACCGGATTAATA	7699
Db	8675	TCTGAGAAATAGTGTATCGGGGAGCGAGTTGCTCTTCCCGCGGCTCAACCGGATTAATA	8734
Qy	7700	CCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGCTTCTTCGGGGCGAA	7759
Db	8735	CCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGCTTCTTCGGGGCGAA	8794
Qy	7760	AACTCTCAAGGATCTTACCCTGTTGAGTTCAGATTCGATTCGATTCGATTCGATTCGATTC	7819
Db	8795	AACTCTCAAGGATCTTACCCTGTTGAGTTCAGATTCGATTCGATTCGATTCGATTCGATTC	8854
Qy	7820	ACTGATCTTACGATCTTTTACTTTTACCGAGCTTCTGCTGAGCAAAAACAGGAAGGC	7879
Db	8855	ACTGATCTTACGATCTTTTACTTTTACCGAGCTTCTGCTGAGCAAAAACAGGAAGGC	8914

```
QY 7880 AAAATGCCGCAAAAAGGGAATAAGGGGACACGGAATGTTGAATCTCATCTCTTCC 7939
DB 8915 AAAATGCCGCAAAAAGGGAATAAGGGGACACGGAATGTTGAATCTCATCTCTTCC 8974
QY 7940 TTTTTCATATATTGAGCATTATATCAGGTTATTTCTCATGAGCGGATACATATTG 7999
DB 8975 TTTTTCATATATTGAGCATTATATCAGGTTATTTCTCATGAGCGGATACATATTG 9034
QY 8000 AATGTATTAGAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAAAGTGCAC 8059
DB 9035 AATGTATTAGAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAAAGTGCAC 9094
QY 8060 CTGAGCTCTAGAACCAATTATATCATGACATTAACCTATAAATAAGCGGTATACGA 8119
DB 9095 CTGAGCTCTAGAACCAATTATATCATGACATTAACCTATAAATAAGCGGTATACGA 9154
QY 8120 GGGCCCTTCGCTTCAAGAAATTAATCTCATGTTTGACAGCTTATCATGATAGCTGAC 8179
DB 9155 GGGCCCTTCGCTTCAAGAAATTAATCTCATGTTTGACAGCTTATCATGATAGCTGAC 9214
QY 8180 TCAATGTTGGTATTGTAATAGACGACAGATCGGAACACTGGAATAAACAAGTTATT 8239
DB 9215 TCAATGTTGGTATTGTAATAGACGACAGATCGGAACACTGGAATAAACAAGTTATT 9274
QY 8240 CG 8241
DB 9275 CG 9276

RESULT 14
ADB97562
ID ADB97562 standard; DNA; 9276 BP.
XX
AC ADB97562;
XX
DT 04-DEC-2003 (first entry)
XX
DE Pichia pastoris vector pPIC9K, SEQ ID NO:13.
XX
KW Type I transmembrane serine protease 25; MTSP25; serine protease;
KW protease cleavage activation; diagnostic marker; neoplastic disorder;
KW cancer; breast; cervix; prostate; lung; ovary; cytosolic; gene therapy;
KW drug screening; tumour progression; monitoring; Pichia pastoris; pPIC9K;
KW cyclic; circular; ds.
XX
OS Synthetic.
XX
FN WO2003031585-A2.
XX
PD 17-APR-2003.
XX
PF 08-OCT-2002; 2002WO-US032417.
XX
PR 09-OCT-2001; 2001US-0328530P.
XX
PA (CORV-) CORVAS INT INC.
XX
FI Madison, EL, Yeh J;
XX
DR WPI; 2003-393442/37.
XX
PT New purified single- or two-chain polypeptide, useful for diagnosing,
PT preventing or treating cancer (e.g. colon cancer), comprises a protease
PT domain of a type-I membrane-type serine protease 25 or its catalytically
PT active portion.
XX
PS Example 2; Page 187-190; 97pp; English.
XX
CC The invention relates to human type I transmembrane serine protease 25
CC (MTSP25; ADB97565) and polypeptides derived from it (e.g., ADB97555). The
CC MTSP25 gene is located on chromosome 12. MTSP25 is a serine protease
CC which is activated by protease cleavage to yield a two-chain protease
```

```
CC comprising an A chain and a B chain linked by a disulphide bond. MTSP25
CC is expressed or is active in tumour cells, and can therefore be used as a
CC diagnostic marker for certain cancers. The invention also encompasses
CC nucleic acids encoding an MTSP25 polypeptide (ADB97564, ADB97571,
CC ADB97554); nucleic acid vectors and host cells comprising an MTSP25
CC polynucleotide; a MTSP25 knockout animal; and an antibody specific for
CC either the single chain (zymogen) or two-chain (activated) form of
CC MTSP25. MTSP25 polypeptides are useful in diagnosing, preventing or
CC treating neoplastic diseases, such as cancer of the breast, cervix,
CC prostate, lung, ovary or colon. MTSP25 polypeptides may also be used for
CC identifying compounds that modulate the protease activity of the
CC polypeptide and for monitoring tumour progression and/or therapeutic
CC effectiveness. The present sequence represents the Pichia pastoris
CC vector pPIC9K. MTSP25 nucleic acid sequences were cloned into a
CC derivative of this vector in an example from the invention
XX
SQ Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;
Query Match 79.6%; Score 6561.6; DB 9; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;
QY 1 AGATCTAATCAATCAAGAGCAAGAGGTGAATGAACCTTTTGGCATCCGATCCACAG 60
DB 1 AGATCTAATCAATCAAGAGCAAGAGGTGAATGAACCTTTTGGCATCCGATCCACAG 60
QY 61 GTCCATTCTCACACATAAGTGCCAAACGACAGGAGGGGATACACTAGCAGCAGACCGT 120
DB 61 GTCCATTCTCACACATAAGTGCCAAACGACAGGAGGGGATACACTAGCAGCAGACCGT 120
QY 121 TCGAAACGAGGACCTCCACTCTCTTCTCTCAACACCCACTTTTGGCATCGAAAACC 180
DB 121 TCGAAACGAGGACCTCCACTCTCTTCTCTCAACACCCACTTTTGGCATCGAAAACC 180
QY 181 AGCCAGATTATGGGCTTGATTGGAGCTCGCTATTCCTCAATTCCTTATTAGGCTACTA 240
DB 181 AGCCAGATTATGGGCTTGATTGGAGCTCGCTATTCCTCAATTCCTTATTAGGCTACTA 240
QY 241 ACACCATGACATTTATTAGCTCTGTATCTCTGCCCCCTCGCGAGGTTCACTGTTGTTTA 300
DB 241 ACACCATGACATTTATTAGCTCTGTATCTCTGCCCCCTCGCGAGGTTCACTGTTGTTTA 300
QY 301 TTTCCGAATGCACAGCTCCGATTACACCGACATCACTCCAGATGAGGGCTTTCTG 360
DB 301 TTTCCGAATGCACAGCTCCGATTACACCGACATCACTCCAGATGAGGGCTTTCTG 360
QY 361 AGTGTGGGGTCAAAATAGTTTTCATGTTTCCCAATAGGCCCAAAACTGACAGTTTAAACGCT 420
DB 361 AGTGTGGGGTCAAAATAGTTTTCATGTTTCCCAATAGGCCCAAAACTGACAGTTTAAACGCT 420
QY 421 GTCTTGGAACTTAATATGACAAAACCGGTGATCTATCCAAAGATGAATTAAGTTGGTTCG 480
DB 421 GTCTTGGAACTTAATATGACAAAACCGGTGATCTATCCAAAGATGAATTAAGTTGGTTCG 480
QY 481 TTGAAATGCTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCATACCGTTTGT 540
DB 481 TTGAAATGCTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCATACCGTTTGT 540
QY 541 CTTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGTTAGCCAGTCT 600
DB 541 CTTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGTTAGCCAGTCT 600
QY 601 CTTCTATCGCTTCTGAACCCCGGTGACCTGTCGGAACGCAATGGGAAACACCCGCT 660
DB 601 CTTCTATCGCTTCTGAACCCCGGTGACCTGTCGGAACGCAATGGGAAACACCCGCT 660
QY 661 TTTTCGATGATTATGATTGTTCTCCATTTGCTTCCAAAGATTTCTGTTGGGAATACT 720
DB 661 TTTTCGATGATTATGATTGTTCTCCATTTGCTTCCAAAGATTTCTGTTGGGAATACT 720
QY 721 GCTGATACCTAAAGTTTCATGATCAAAATTAATCTTAACCCCTACTTGCAGCAAT 780
DB 721 GCTGATACCTAAAGTTTCATGATCAAAATTAATCTTAACCCCTACTTGCAGCAAT 780
```





QY	4829	-----	4828
Db	4955	AAACTGCAATTTATTCATATCAGGATTTATCAATACCATAATTTTGGAAAAGCCGTTTCTG	5014
QY	4829	-----	4828
Db	5015	TAATGAAGGAGAAAACCTCACCAGGCGAGTTCCATAGGATGGCAAGATCTTGATCGGTC	5074
QY	4829	-----	4828
Db	5075	TGCGATTCGGAATCGTCCACATCAATACACCTATTAAATTCCTCGTCAAAAATAAG	5134
QY	4829	-----	4828
Db	5135	GTATCAAGTGAGAAATCACCATGATGACGACTGAATCCGGTGAGAAATGGCAAAAGCTT	5194
QY	4829	-----	4828
Db	5195	ATGCATTTCTTCCAGACTTGTTCACAGGCGAGCCATTACGCTCGTCATCAAAATCACT	5254
QY	4829	-----	4828
Db	5255	CGCATCAACCAACCGTTATTCATTCGTGATTGGCCCTGAGCGAGACGAAATACGCGATC	5314
QY	4829	-----	4828
Db	5315	GCTGTAAAAGGACAATTAACAACAGGAATCGAATGCAACCGCGCGAGAACATGCGCAG	5374
QY	4829	-----	4828
Db	5375	CGCATCAACAATTTTCACTGGAATCAGGATATTCCTTAATACCTGGATGCTGTTT	5434
QY	4829	-----	4828
Db	5435	CCCGGGATCGAGTGGTGAATACCATGCATCATCAGGAGTACGGAATAAATGCTTGAT	5494
QY	4829	-----	4828
Db	5495	GGTCGGAAGAGCAATAATTCGGTCAGCCAGTTTGTAGTCGACCATCTCATCTGTAAACATC	5554
QY	4829	-----	4828
Db	5555	ATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCAT	5614
QY	4829	-----	4828
Db	5615	CAATCGATAGATTGTGCGCACCTGATTCGCCGACATATTCGAGAGCCCATTTATACCCATA	5674
QY	4829	-----	4828
Db	5675	TAAATCAGCATCCATGTTTGGAAATTAATCGCGCCTCGAGCAGACGTTTCCCGTTGAAT	5734
QY	4829	-----	4828
Db	5735	ATGGCTCATAACACCCCTGTTATTTACTGTTTATGTAAGCAGACAGTTTATTTGTTCAATGA	5794
QY	4829	-----	4828
Db	5795	TGATATATTTTATCTTGTGCAATGTAAATCATCAGAGATTTTGAGACAAAGTGCCTTTC	5854
QY	4829	-----	4879
Db	5855	CCCCCCCCCTCAGTGGGATCACCGCGCCACAGGTGGGTTGCTGGCGCCTATAT	5914
QY	4880	CGCGGACATCACCGATGGGAAGATCGGCTCGCACTTCGGGCTCATGAGCGCTGTTT	4939
Db	5915	CGCGGACATCACCGATGGGAAGATCGGCTCGCACTTCGGGCTCATGAGCGCTGTTT	5974
QY	4940	CGCGGATGATGCTGGCAGGCCCCCGTGGCCGGGACCTGTTGGGCGCCATCTCTTGCA	4999
Db	5975	CGCGGATGATGCTGGCAGGCCCCCGTGGCCGGGACCTGTTGGGCGCCATCTCTTGCA	6034

QY	5000	TGCACCATTCCTTGGCGGGCGGTCTCAACGGCCTCAACCTACTACTGGGCTGCTTCT	5059
Db	6035	TGCACCATTCCTTGGCGGGCGGTCTCAACGGCCTCAACCTACTACTGGGCTGCTTCT	6094
QY	5060	AATGCAAGAGTCGCATTAAGGAGAGCGTCAATATCTATGATTTGGAAGTATCGGAATGT	5119
Db	6095	AATGCAAGAGTCGCATTAAGGAGAGCGTCAATATCTATGATTTGGAAGTATCGGAATGT	6154
QY	5120	GATACCCGCAATCTTCAGTGTCTTGAAGTCTCTCTATCAGATTAATGCGCCCAATAGCAAC	5179
Db	6155	GATACCCGCAATCTTCAGTGTCTTGAAGTCTCTCTATCAGATTAATGCGCCCAATAGCAAC	6214
QY	5180	CGGAGGAGAGATTTTCATGTAATTTCTCTGACTTTTGGTCAATCAGTAGACTCGAACTG	5239
Db	6215	CGGAGGAGAGATTTTCATGTAATTTCTCTGACTTTTGGTCAATCAGTAGACTCGAACTG	6274
QY	5240	TGAGACTATCTCGGTTATCAGCAGAGAAATGCTCTTTGGAGAGAGTAATAGAACTCC	5299
Db	6275	TGAGACTATCTCGGTTATCAGCAGAGAAATGCTCTTTGGAGAGAGTAATAGAACTCC	6334
QY	5300	ACCAATAAGAAATCCTTGTATCAGAAACAACTCTTGTTCGAACTTTTTCGGTGCC	5359
Db	6335	ACCAATAAGAAATCCTTGTATCAGAAACAACTCTTGTTCGAACTTTTTCGGTGCC	6394
QY	5360	TTGAACTATAAAATGTAGAGTGGATATGTCGGGTAGGAATGGAGCGGGCAATGCTTACC	5419
Db	6395	TTGAACTATAAAATGTAGAGTGGATATGTCGGGTAGGAATGGAGCGGGCAATGCTTACC	6454
QY	5420	TTCTGACACTTCAAGAGTATGTAGAGTTGTAGACTGATGCCAACTTCAGTGACAAAC	5479
Db	6455	TTCTGACACTTCAAGAGTATGTAGAGTTGTAGACTGATGCCAACTTCAGTGACAAAC	6514
QY	5480	GTTGCTATTTCGTTCAAAACCAATCCGAATCCAGAGAAATCAAAAGTTGTTGCTACTATT	5539
Db	6515	GTTGCTATTTCGTTCAAAACCAATCCGAATCCAGAGAAATCAAAAGTTGTTGCTACTATT	6574
QY	5540	GATCCAAAGCAGTCGGTCTTGAACCTGACATAGTGTGCTGCTGTTTGGAGTCACTTT	5599
Db	6575	GATCCAAAGCAGTCGGTCTTGAACCTGACATAGTGTGCTGCTGTTTGGAGTCACTTT	6634
QY	5600	TGTATGAATAAATCTAGTCTTTCATCTAAATATCTTGACGAGCCAGGCGATAAATACC	5659
Db	6635	TGTATGAATAAATCTAGTCTTTCATCTAAATATCTTGACGAGCCAGGCGATAAATACC	6694
QY	5660	CAAACTAAAACCTTTTAAAACGTTTAAAAGCAAGTATGTCGCTGTATTTAAACCCC	5719
Db	6695	CAAACTAAAACCTTTTAAAACGTTTAAAAGCAAGTATGTCGCTGTATTTAAACCCC	6754
QY	5720	AAATCAGCTGATGATCTCTCATCAACTTGAGGGGCACTATCTGTTTTAGAGAAAT	5779
Db	6755	AAATCAGCTGATGATCTCTCATCAACTTGAGGGGCACTATCTGTTTTAGAGAAAT	6814
QY	5780	TTGCGGAGATCGGATATCGAGAAAAAGGTACGCTGATTTTAAACGTTGAAATTTATCTCAA	5839
Db	6815	TTGCGGAGATCGGATATCGAGAAAAAGGTACGCTGATTTTAAACGTTGAAATTTATCTCAA	6874
QY	5840	GATCTCTGCTCGCGCTTTTCGGTGTATGACGTTGAAACCTCTGACATGCGAGCTCCG	5899
Db	6875	GATCTCTGCTCGCGCTTTTCGGTGTATGACGTTGAAACCTCTGACATGCGAGCTCCG	6934
QY	5900	GAGACGTCACAGCTTCTCTGTAAGCGGATGCGGGAGCAGACAGCCCGTCAGGGCGCG	5959
Db	6935	GAGACGTCACAGCTTCTCTGTAAGCGGATGCGGGAGCAGACAGCCCGTCAGGGCGCG	6994
QY	5960	TCAGCGGCTGTGGCGGTGTGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGA	6019
Db	6995	TCAGCGGCTGTGGCGGTGTGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGA	7054
QY	6020	GTGTATCTGCTTAACTATGCGCATCAGACGAGTTGTACTGAGAGTGCACCATATGC	6079
Db	7055	GTGTATCTGCTTAACTATGCGCATCAGACGAGTTGTACTGAGAGTGCACCATATGC	7114
QY	6080	GCTGTGAAATACCGCACAGATGCGTAAAGGAAAAATACCGCATCAGGCGCTTCTTCCGCTT	6139





RESULT 15  
AAD09083  
ID AAD09083 standard; DNA; 9776 BP.  
XX AC  
XX AAD09083;  
XX  
DT 04-SEP-2001 (first entry)  
XX  
XX  
DE PUB55 expression plasmid for Sonic hedgehog protein in Pichia pastoris.  
XX  
XX Human; hedgehog protein; nontropic; neuroprotective; anticonvulsant;  
KW cytosatic; therapy; Alzheimer's disease; Parkinson's disease; injury;  
KW Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis;  
KW nervous system aging; neurodegenerative disease; immunological disease;  
KW malignant glioma; medulloblastoma; neuroectodermal tumour; cancer;  
KW extracellular signalling protein; plasmid PUB55; ds.  
XX  
XX Homo sapiens.  
OS Unidentified.  
OS Chimeric.  
XX  
XX WO200134654-A1.  
XX  
XX 17-MAY-2001.  
PD  
XX  
XX 02-NOV-2000; 200WO-US030405.  
FF  
XX  
XX 05-NOV-1999; 99US-0164025P.  
PR  
XX  
XX (BIOJ ) BIOGEN INC.  
PA  
XX  
XX Strauch K;  
PI  
XX  
XX WPI; 2001-329075/34.  
DR  
XX  
XX Novel isolated hedgehog fusion polypeptide useful for treating  
PT neurological conditions such as Alzheimer's disease, Parkinson's disease,  
PT Huntington's chorea, amyotrophic lateral sclerosis, and multiple  
PT sclerosis.  
XX  
XX Example 1; Page 62-67; 178pp; English.  
PS  
XX  
XX The present invention relates to hedgehog fusion proteins. Hedgehog  
CC proteins are a family of extracellular signalling proteins that regulate  
CC various aspects of embryonic development both in vertebrates and in  
CC invertebrates. Hedgehog fusion protein is useful for the prophylaxis or  
CC treatment of any condition or disease state for which a hedgehog or  
CC patched protein constituent is efficacious and in the diagnosis of  
CC constituents or conditions of disease states in biological systems or  
CC specimens and for diagnostic purposes in non-physiological systems.  
CC Hedgehog fusion protein is useful for treating neurological conditions  
CC due to injury, aging of nervous system, including Alzheimer's disease,  
CC chronic neurodegenerative diseases of the nervous system, including  
CC Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis  
CC and chronic immunological diseases of nervous system including multiple  
CC sclerosis and malignant gliomas, medulloblastomas, neuroectodermal  
CC tumours and to specifically target medical therapies against cancers and  
CC tumours which express the receptor for the protein. The present sequence  
CC is PUB55 expression plasmid for sonic hedgehog protein in Pichia  
CC pastoris. The plasmid contains the DNA encoding N-terminal domain of  
CC human sonic hedgehog with the alpha factor prepro region as the secretion  
CC signal  
XX  
XX Sequence 9776 BP; 2608 A; 2305 C; 2263 G; 2600 T; 0 U; 0 Other;  
SQ  
  
Query Match 77.5%; Score 6383.8; DB 4; Length 9776;  
Best Local Similarity 82.9%; Fred. No. 0;  
Matches 8101; Conservative 0; Mismatches 137; Indels 1535; Gaps 10;  
  
2 GATCTTAACATCCAAAGACCAAGGTTGTAATGAACCTTTTCCATCCGACATCCACAGG 61

QY 1142 TTATAAATCTACTATTGCGAGCATCTGCTAAAGAAAGAGGGTATCTCTCGAAGAAA 1201  
DB 1141 TTTATTAATCTACTATTGCGAGCATCTGCTAAAGAAAGAGGGTATCTCTCGAAGAAA 1200  
QY 1202 GAGAGGCTGAAGCCGAGG- 1219  
DB 1201 GATCGGACCGGAGGCGGTTTCGGAAGAGAGGACCCCAAAAAGCTGACCCCTTTAG 1260  
QY 1220 - 1220 - - - - - CCGAGCCAGGCCCCAGGCCAG 1242  
DB 1261 CCTACAAGAGTTTATCCCAATGTGCGGAGAAAGACCTTAGGCGCGAGGAGGTATG 1320  
QY 1243 GAAGAAACAT- 1252  
DB 1321 AAGGAGATCTCCAGAACTCCGAGCGATTTAAGGAACCTCACCCCAATTACAACCCCG 1380  
QY 1253 - - - - - GTAATCATGCGCCCGCGGAGAGGATTAATTTGTGCTTCCCGGTGTC 1299  
DB 1381 ACATCATATTTAAGGATGAAGAAACACCGGAGCGGACGCTGATGACTCAGAGGTGA 1440  
QY 1300 AC- - - - - CGCCAGCAGTGACGAGAGAGGTTGCTGTTT 1334  
DB 1441 AGGCAAGTTGAACGCTTTGGCCATCTCGGTGATGAACCCAGTGCCAGGAGTGAACCTGC 1500  
QY 1335 TGATGACAGTGTCCGGGAGTTCCCGTG- 1362  
DB 1501 GGGTGACCGAGGCTGGGACGAAGATGGCCACCACCTCAGAGGAGTCTCTGCACTACGAGG 1560  
QY 1363 - - - - - 1362  
DB 1561 GCCGGCAGTGACATCACACGCTCTGACCGGACCGCAGCAAGTACGGCATGCTGCGCC 1620  
QY 1363 - - - - - TGTTCACCCCATGCGCCATCGAGAACACTCAAGAAAGAAAT 1405  
DB 1621 GCCTGGCGGTGAGGCGCGCTTCGACTGGGTGTACTAGAGTCAAGGACACATATCCACT 1680  
QY 1406 GTCCCTCTCACTA- - - - - GTGGGTAGATTCCTAGG- - - - - GCGGCGCGA 1449  
DB 1681 GCTCGGTGAAGACAGAGAACTCGGTGGCGGCAATCGGAGGCTGATTCGCGCGCGCGA 1740  
QY 1450 ATTAATTCGCTTAGACATGACTGTCTCAGTTCAAGTTGGGCACTTACGAGAAAGACCG 1509  
DB 1741 ATTAATTCGCTTAGACATGACTGTCTCAGTTCAAGTTGGGCACTTACGAGAAAGACCG 1800  
QY 1510 GTCTTGCTAGATTTCTAATCAAGAGGATGTGAAATGCCATTTGCTGAGAGATGCGAGGCT 1569  
DB 1801 GTCTTGCTAGATTTCTAATCAAGAGGATGTGAAATGCCATTTGCTGAGAGATGCGAGGCT 1860  
QY 1570 TCATTTTGTACTTTTATTTGTAACCTATATAGTATAGGATTTTGTGTCATTTTG 1629  
DB 1861 TCATTTTGTACTTTTATTTGTAACCTATATAGTATAGGATTTTGTGTCATTTTG 1920  
QY 1630 TTTCTTCTGACGAGTGTGCTGATCAGCTTATCTGCGAGCTGATGAATATCTTG 1689  
DB 1921 TTTCTTCTGACGAGTGTGCTGATCAGCTTATCTGCGAGCTGATGAATATCTTG 1980  
QY 1690 GTAGGGTTTGGGAAATCATTCAGTTGTGATGTTTTCTTGTTATTTCCCACTCTCTT 1749  
DB 1981 GTAGGGTTTGGGAAATCATTCAGTTGTGATGTTTTCTTGTTATTTCCCACTCTCTT 2040  
QY 1750 CAGAGTACAGAAGATTTAAGTGAAGTTTCTGTTGCAAGCTTATCGATAAGCTTTAATG 1809  
DB 2041 CAGAGTACAGAAGATTTAAGTGAAGTTTCTGTTGCAAGCTTATCGATAAGCTTTAATG 2100  
QY 1810 CGGTAGTTTATCACAGTTAAATTTGTAACGAGTCAGGACCGGTGATGAATCTAACAA 1869  
DB 2101 CGGTAGTTTATCACAGTTAAATTTGTAACGAGTCAGGACCGGTGATGAATCTAACAA 2160  
QY 1870 TGCCTCATCTGCTACCTCGGACCGCTCACCCCTGGATGCTGTAGGCACTAGGCTTGTTAT 1929  
DB 2161 TGGCTCATCTGCTACCTCGGACCGCTCACCCCTGGATGCTGTAGGCACTAGGCTTGTTAT 2220

QY 1930 GCGGTACTTGGCGGCTCTTTCGGGATATCGTCCATTCGACAGCATGCGCAGTCACTA 1989  
DB 2221 GCGGTACTTGGCGGCTCTTTCGGGATATCGTCCATTCGACAGCATGCGCAGTCACTA 2280  
QY 1990 TGGCGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGACCCGTTCTCGGAGC 2049  
DB 2281 TGGCGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGACCCGTTCTCGGAGC 2340  
QY 2050 ACTGTCCGACCGCTTTGGCGCGCCCGCCAGTCTCTGCTCGCTTCGTTCTGAGCCACTAT 2109  
DB 2341 ACTGTCCGACCGCTTTGGCGCGCCCGCCAGTCTCTGCTCGCTTCGTTCTGAGCCACTAT 2400  
QY 2110 CGATPACGGGATCATGCGGACACACCCGTCCTGTGATCTATCGAATCTAAATGTAAGT 2169  
DB 2401 CGACTPACGGGATCATGCGGACACACCCGTCCTGTGATCTATCGAATCTAAATGTAAGT 2460  
QY 2170 TAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGAACCTTAAACAGATTGC 2229  
DB 2461 TAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGAACCTTAAACAGATTGC 2520  
QY 2230 GGTGAGCATCTAGACCTTTCAACAGCGCCAGATCCATCACTGCTTGGCCAAATATGTTCA 2289  
DB 2521 GGTGAGCATCTAGACCTTTCAACAGCGCCAGATCCATCACTGCTTGGCCAAATATGTTCA 2580  
QY 2290 GTCCCTCAGGAGTACGTCCTTGTGAAGTATGAACCTTCTGGAAGTTGCACTGTTAACTC 2349  
DB 2581 GTCCCTCAGGAGTACGTCCTTGTGAAGTATGAACCTTCTGGAAGTTGCACTGTTAACTC 2640  
QY 2350 CGCTGTATTTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGTTACCGGAGGAGTAAT 2409  
DB 2641 CGCTGTATTTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGTTACCGGAGGAGTAAT 2700  
QY 2410 CTCACAATCTCTGAGAGTAGGACCAACAAACAGACAGATCCAGCGTGTGTTACTTGAT 2469  
DB 2701 CTCACAATCTCTGAGAGTAGGACCAACAAACAGACAGATCCAGCGTGTGTTACTTGAT 2760  
QY 2470 CAACATAAGAAAGACATTTCTCGATTTTCAGAGTCAAGTGTTCAGGAGCGTACTGATTGG 2529  
DB 2761 CAACATAAGAAAGACATTTCTCGATTTTCAGAGTCAAGTGTTCAGGAGCGTACTGATTGG 2820  
QY 2530 ACATTTCCAAAGCCTGCTGATGTTGCAACCGATAGGTTGTAGTGTGCAATACACT 2589  
DB 2821 ACATTTCCAAAGCCTGCTGATGTTGCAACCGATAGGTTGTAGTGTGCAATACACT 2880  
QY 2590 TCGGTACAAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTGTGAACAGCATCTTCAATTC 2649  
DB 2881 TCGGTACAAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTGTGAACAGCATCTTCAATTC 2940  
QY 2650 TGGCAAGCTCTCTGCTCATATCGACAGCAACAGAAATCACCTGGGAATCAATACCAT 2709  
DB 2941 TGGCAAGCTCTCTGCTCATATCGACAGCAACAGAAATCACCTGGGAATCAATACCAT 3000  
QY 2710 GTTCAGCTTGAGACAGAGGCTGAGGCAACGAATCTGGATCAGCGTATTTATCAGCAA 2769  
DB 3001 GTTCAGCTTGAG- CAGAAAGCTGAGGCAACGAATCTGGATCAGCGTATTTATCAGCAA 3059  
QY 2770 TAACTAGAACTTCAGAAAGCCCGCAGGAGCATGTCAATACTACAGGCGCTGATGTGAT 2829  
DB 3060 TAACTAGAACTTCAGAAAGCCCGCAGGAGCATGTCAATACTACAGGCGCTGATGTGAT 3119  
QY 2830 TTTGAACCATCATCTTGGCAGGATTAACGAATCGTTTCTTGGACCAATATTTGTCAC 2889  
DB 3120 TTTGAACCATCATCTTGGCAGGATTAACGAATCGTTTCTTGGACCAATATTTGTCAC 3179  
QY 2890 ACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTTGGCGGCTCTCTGCTA 2949  
DB 3180 ACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTTGGCGGCTCTCTGCTA 3239  
QY 2950 GCAAGATACATTTAGCACCAACCTTGTGGGCAACGTAGATGACTTCTGGGGTAGGGTAC 3009  
DB 3240 GCAAGATACATTTAGCACCAACCTTGTGGGCAACGTAGATGACTTCTGGGGTAGGGTAC 3299  
QY 3010 CATCTCTTCTTAGGTGGAGATGCAAAAAACAATTTCTTTCACACCAAGCACTTTTGGCAGAA 3069

Db 3300 CATCTCTTAGTGGAGATCAAAAACAAATTTCTTGGACACGCAACTTTGGCAGAA 3359  
Qy 3070 CACCAGCATCAGGAGTGAAGGCAGAAATTCGGTTCCACACGAGATATAGAGGCCAA 3129  
Db 3360 CACCAGCATCAGGAGTGAAGGCAGAAATTCGGTTCCACACGAGATATAGAGGCCAA 3419  
Qy 3130 CTTTCTCAATAGTCTTCAAAAAGAGAGCAGACTACACAGGGCAAGTCTCAACTTGA 3189  
Db 3420 CTTTCTCAATAGTCTTCAAAAAGAGAGCAGACTACACAGGGCAAGTCTCAACTTGA 3479  
Qy 3190 ACGTCTCGTTAGTGGATTCATGGAATTTCTGACGTTATCTATAGAGATCAATGG 3249  
Db 3480 ACGTCTCGTTAGTGGATTCATGGAATTTCTGACGTTATCTATAGAGATCAATGG 3539  
Qy 3250 CTTCTTAACGTTATCTGGCAATTCGATAAGTCTCTGGGAAAGGAGCTTCTAACACAG 3309  
Db 3540 CTTCTTAACGTTATCTGGCAATTCGATAAGTCTCTGGGAAAGGAGCTTCTAACACAG 3599  
Qy 3310 GTGTCTCAAGGAGCTCCATCAACCTTGGCAGTTAGTCTTAAAGGGCTTTGTCACAT 3369  
Db 3600 GTGTCTCAAGGAGCTCCATCAACCTTGGCAGTTAGTCTTAAAGGGCTTTGTCACAT 3659  
Qy 3370 TTTGAGCAACATTTGCGCAATTTGGTTGACTAATCCATAATCTGTTCGGTTTCTGGA 3429  
Db 3660 TTTGAGCAACATTTGCGCAATTTGGTTGACTAATCCATAATCTGTTCGGTTTCTGGA 3719  
Qy 3430 TAGGACGAGGAGCATCTTCAATTTCTTGTGAGGAGCCCTTAAAGCGTCAATTTGTC 3489  
Db 3720 TAGGACGAGGAGCATCTTCAATTTCTTGTGAGGAGCCCTTAAAGCGTCAATTTGTC 3779  
Qy 3490 ACAATTCATAGGACCTTCAGAGGAGCTTCTTGTAGTTTGGATTCCTTTAGTGTGTT 3549  
Db 3780 ACAATTCATAGGACCTTCAGAGGAGCTTCTTGTAGTTTGGATTCCTTTAGTGTGTT 3839  
Qy 3550 CTTTGGTGTATCTGGTGTGATCTCTCTTTCCTTCTAGTGACCTTTAGGACCTTCATAT 3609  
Db 3840 CTTTGGTGTATCTGGTGTGATCTCTCTTTCCTTCTAGTGACCTTTAGGACCTTCATAT 3899  
Qy 3610 CAGGTTTCTCTCACCTCTGTCACAGTCACACCGTACTTGGCACATCTAATATGCAA 3669  
Db 3900 CAGGTTTCTCTCACCTCTGTCACAGTCACACCGTACTTGGCACATCTAATATGCAA 3959  
Qy 3670 AATAAATAAGTCAGCACATTCACAGGCTATCTCTCTGATTTAGCTTCTGCAAGTT 3729  
Db 3960 AATAAATAAGTCAGCACATTCACAGGCTATCTCTCTGATTTAGCTTCTGCAAGTT 4019  
Qy 3730 CATAGCTTCTCTCAATTTTAGCGTTC - AACAAAATTCGTCTGCAATTAACCGTTG 3788  
Db 4020 CATAGCTTCTCTCAATTTTAGCGTTC - AACAAAATTCGTCTGCAATTAACCGTTG 4079  
Qy 3789 GTATAGAACCTTCTGAGCATTTCTTACGATCCACAGGTGGCTTCCATGGCTCTA 3848  
Db 4080 GTATAGAACCTTCTGAGCATTTCTTACGATCCACAGGT - GCTTCCATGGCTCTA 4138  
Qy 3849 AGACCTTTGATTTGGCCAAAACAGGAAGTGGTTC - AACAAAATTCGTCTGCAATTAACCGTTG 3908  
Db 4139 AGACCTTTGATTTGGCCAAAACAGGAAGTGGTTC - AACAAAATTCGTCTGCAATTAACCGTTG 4198  
Qy 3909 TGTTCACACAAAATTCAGCAGTCTCCATCAATTCGAATCCAGCAACTTT 3968  
Db 4199 TGTTCACACAAAATTCAGCAGTCTCCATCAATTCGAATCCAGCAACTTT 4258  
Qy 3969 TCAGTTGCTCCAGATGTAGACCTTTATACCAAAACCGTACGACGAGATTGGTAGCT 4028  
Db 4259 TCAGTTGCTCCAGATGTAGACCTTTATACCAAAACCGTACGACGAGATTGGTAGCT 4318  
Qy 4029 CCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAAC 4088  
Db 4319 CCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAAC 4378  
Qy 4089 GAGTAATAGAGAGTCAGCCACCAAGTGTGAATAGACCATCGGGGGTCTAGTAGTC 4148

Db 4379 GAGTAATAGAGAGTCAGCCACCAAGAGTAGTAATAGACCATCGGGCGGTCTAGTAGTC 4438  
Qy 4149 AAAGACGCCCAACAAAATTTTCTGACAGGAACTTTTTCACATCTTTCAGAAAGTTCTGAT 4208  
Db 4439 AAAGACGCCCAACAAAATTTTCTGACAGGAACTTTTTCACATCTTTCAGAAAGTTCTGAT 4498  
Qy 4209 TCAGTAGTCAATTCGCGAGCATCAATATATGGGATTTATACAGAGCAACAGTGGAGTC 4268  
Db 4499 TCAGTAGTCAATTCGCGAGCATCAATATATGGGATTTATACAGAGCAACAGTGGAGTC 4558  
Qy 4269 ACATCTACCAACTTTTTCGGTCTCAGAAAAGCATAAACAGTTTCTACTACCGCATTTAGTG 4328  
Db 4559 ACATCTACCAACTTTTTCGGTCTCAGAAAAGCATAAACAGTTTCTACTACCGCATTTAGTG 4618  
Qy 4329 AAACTTTTCAATTCGCCAGTGGAGAAAAGGACAGCATCTAGCATTTAGCGGC 4388  
Db 4619 AAACTTTTCAATTCGCCAGTGGAGAAAAGGACAGCATCTAGCATTTAGCGGC 4678  
Qy 4389 AAGGATGCAACTTTTATCAACAGGCTCTATAGATAACCTAGCGCTGGGATCATCTT 4448  
Db 4679 AAGGATGCAACTTTTATCAACAGGCTCTATAGATAACCTAGCGCTGGGATCATCTT 4738  
Qy 4449 TGGACAACTCTTTTCTGCCAAATCTAGGTCAAAATCAGTTTCAATTTATGATCACTTATGTC 4508  
Db 4739 TGGACAACTCTTTTCTGCCAAATCTAGGTCAAAATCAGTTTCAATTTATGTC 4798  
Qy 4509 AACTTGAGCAAGTTTGTGATCAGCTCTCAAAATTTGCTCTCTGTAACGATGACTCAACT 4568  
Db 4799 AACTTGAGCAAGTTTGTGATCAGCTCTCAAAATTTGCTCTCTGTAACGATGACTCAACT 4858  
Qy 4569 TGCACATTAACCTTGAAGCTCAGTCGATTCAGTGAACCTTCAATTCAGTTGTCAGCTCGTCA 4628  
Db 4859 TGCACATTAACCTTGAAGCTCAGTCGATTCAGTGAACCTTCAATTCAGTTGTCAGCTCGTCA 4918  
Qy 4629 GCAGCATAGGAAACACGGCTTTTCTACCAAACTCAAGAAATTTTCAAACTCTGCAACA 4688  
Db 4919 GCAGCATAGGAAACACGGCTTTTCTACCAAACTCAAGAAATTTTCAAACTCTGCAACA 4978  
Qy 4689 CTTTGCCTTGCAGTAGCAGGAGATTCATCTTGAAGTGGACAGTGGAGTGTAGTCT 4748  
Db 4979 CTTTGCCTTGCAGTAGCAGGAGATTCATCTTGAAGTGGACAGTGGAGTGTAGTCT 5038  
Qy 4749 TGAGAAATTTCTGAAGCGTATTTTATATCAGTGTAGTCAATCAGGAGATTCCTCTAC 4808  
Db 5039 TGAGAAATTTCTGAAGCGTATTTTATATCAGTGTAGTCAATCAGGAGATTCCTCTAC 5098  
Qy 4809 GCCGAGCATCTGTGCGGACCTTGCAGTTCGATCAGGATCCGCGCCACAGTGGGTGCT 4868  
Db 5099 GCCGAGCATCTGTGCGGACCTTGCAGTTCGATCAGGATCCGCGCCACAGTGGGTGCT 5158  
Qy 4869 GCGGCTTATATCGCGACATCACCGATGGGAGATCGGCTCGCCACTTTGGGCTCATG 4928  
Db 5159 GCGGCTTATATCGCGACATCACCGATGGGAGATCGGCTCGCCACTTTGGGCTCATG 5218  
Qy 4929 AGCGTTGTTTTCGGCTGGGTATGTTGGCAGGCGCGTGGCGGGGAGTGTGGGCGCC 4988  
Db 5219 AGCGTTGTTTTCGGCTGGGTATGTTGGCAGGCGCGTGGCGGGGAGTGTGGGCGCC 5278  
Qy 4989 ATCTCCTTG - 4997  
Db 5279 ATCTCCTTGACCTGACGGGGGGGGGAAAGCCACGTTGTCTCTCAAAATCTCTGAT 5338  
Qy 4998 - 4997  
Db 5339 GTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAACTGTCTGTCTACATAA 5398  
Qy 4998 - 4997  
Db 5399 ACAGTAATACAGGGGTGTTATAGCCATATTCAACGGGAAACGCTTTGCTCAAGCGCC 5458  
Qy 4998 - 4997  
Db 5459 GATTAAATCCACATGGATGCTGATTATATGGGTATAAATGGGCTCGCGATAATGTCG 5518

QY 4998 ----- 4997  
Db 5519 GGCATCAGGTGCGACAAATATATCGATTGTATGGGAAGCCGATGCGCCAGAGTTGTTTC 5578  
QY 4998 ----- 4997  
Db 5579 TGAACATGCGCAAGGTAGCGTTGCCAATGATGTACAGATGAGATGGTTCAGACTAAACT 5638  
QY 4998 ----- 4997  
Db 5639 GGCTGACGGAAATTTATGCTCTCCGACCAATCAAGCAATTTATCCGTACTCTCTGATGATG 5698  
QY 4998 ----- 4997  
Db 5699 CATGTTACTCACCACCTGCGATCCCGGAAACAGCANTCCAGGTATTAGAGAATATC 5758  
QY 4998 ----- 4997  
Db 5759 CTGATTCAGGTGAAATAATTTGTGATCGCTGCGAGTTCCTGCGCCGGTTGCAATTCGA 5818  
QY 4998 ----- 4997  
Db 5819 TTCCTGTTGTAATGTCCITTTAAACAGCGATCCGCTATTTGCTCTCGCTCAGCGCAAT 5878  
QY 4998 ----- 4997  
Db 5879 CACGAATGAATAACGGTTTGGTGTATGCGAGTGAATTTGTATGACGAGCTTAATGGCTGGC 5938  
QY 4998 ----- 4997  
Db 5939 CTGTTGAACAAGTCTGGAAGAAATGCAATAGCTTTTGCCATTTCTACCGGATTCAGTCG 5998  
QY 4998 ----- 4997  
Db 5999 TCACTCATGGTGAATTTCTCACTTGATAAACCCTTATTTTGGACGAGGGAATTAATAGTT 6058  
QY 4998 ----- 4997  
Db 6059 GTATTGATGTTGGACGAGTCGGAATCGACAGCATACCAGGATCTTGCCATCCTATGGA 6118  
QY 4998 ----- 4997  
Db 6119 ACTGCTCGGTGAGTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGTTATG 6178  
QY 4998 ----- 4997  
Db 6179 ATAATCCTGATATGAATAAATTCGAGTTTCATTTGATGCTCGATGAGTTTTCATATCAG 6238  
QY 4998 ----- 4997  
Db 6239 AATTGGTTAATTTGTTGTAACACTGCGACAGCATTTAGCTGACTTGACGGGACGGCGGT 6298  
QY 4998 ----- 4997  
Db 6299 TTGTTGAATAAATCGAACTTTTGTGATGAGTTGAGGATCAGATCAGCATCTTCCCGCAA 6358  
QY 4998 ----- 4997  
Db 6359 CGCAGACCGTTCCGTGGCAAGCAAAAGTTCAAAATCACCNAATCGGTCCACCTACACAA 6418  
QY 4998 ----- 4997  
Db 6419 AGCTCTCATCAACCGTGGCTCCCTCACTTTCTGCTGGATGATGGGCGGATTCAGGCCCTG 6478  
QY 4998 ----- 4997  
Db 6479 GTATGATCAACAACACTTTTTCAGGAGGACACCTCAGCGCCGCCCCCCCCCTGCGG 6538  
QY 5009 CTTTGGCGCGGGTGTCTCAACGGCCTCAACTACTACTGGGTGCTTCTTAATGACGGA 5068  
Db 6539 TCCACGCGCGGTGCTCAACGGCCTCAACTACTACTGGGTGCTTCTTAATGACGGA 6598

QY 5069 GTCCGATAAGGAGAGCGTCCAGTATCTATGATTGGAAGTATGGAATGTTGATACCCGC 5128  
Db 6599 GTCCGATAAGGAGAGCGTCCAGTATCTATGATTGGAAGTATGGAATGTTGATACCCGC 6658  
QY 5129 ATTCTTCAGTCTCTGAGTCTCCTATCAGATTATGCCCACTAAGCAACCGGAGGAGG 5188  
Db 6659 ATTCTTCAGTCTCTGAGTCTCCTATCAGATTATGCCCACTAAGCAACCGGAGGAGG 6718  
QY 5189 AGATTTCATGTTAAATTTCTCTGACTTTTGTCTATCAGTAGACTCGAACTGTGAGACTAT 5248  
Db 6719 AGATTTCATGTTAAATTTCTCTGACTTTTGTCTATCAGTAGACTCGAACTGTGAGACTAT 6778  
QY 5249 CTCGGTTATGACAGCAGAAATGTCTCTTCTGGAGACAGTAATCAAGTCCACCACAAATAA 5308  
Db 6779 CTCGGTTATGACAGCAGAAATGTCTCTTCTGGAGACAGTAATCAAGTCCACCACAAATAA 6838  
QY 5309 GAAATCCCTTGTATCAGGAACAAACTTCTTCTTCTGAACTTTTTCGGTGCCTTGAACAT 5368  
Db 6839 GAAATCCCTTGTATCAGGAACAAACTTCTTCTTCTGAACTTTTTCGGTGCCTTGAACAT 6898  
QY 5369 AAAATGTAGATGGAATATGTCGGTAGGAATGGAAGCGGCAAAATGCTTACCTTCGACC 5428  
Db 6899 AAAATGTAGATGGAATATGTCGGTAGGAATGGAAGCGGCAAAATGCTTACCTTCGACC 6958  
QY 5429 TTCAAGAGTATGAGGTTTGTAGACTACTCATCGCAACTTCAGTGAACAAGTTGCTATT 5488  
Db 6959 TTCAAGAGTATGAGGTTTGTAGACTACTCATGTCGCAACTTCAGTGAACAAGTTGCTATT 7018  
QY 5489 TCGTTCAAAACCATTCGGAATCCAGAGAAATCAAAAGTTGTTGTCTACTATGATCCAAAGC 5548  
Db 7019 TCGTTCAAAACCATTCGGAATCCAGAGAAATCAAAAGTTGTTGTCTACTATGATCCAAAGC 7078  
QY 5549 CAGTCCGCTCTGAACTGCAATAGTCTGCTGCTGTTTGGAGTCACTCTTGTGATGAAT 5608  
Db 7079 CAGTCCGCTCTGAACTGCAATAGTCTGCTGCTGTTTGGAGTCACTCTTGTGATGAAT 7138  
QY 5609 AAATCTAGTCTTGAATCTAAATAATCTTTCAGGAGCCAAAGCGGATAAATACCCAAATCTAA 5668  
Db 7139 AAATCTAGTCTTGAATCTAAATAATCTTTCAGGAGCCAAAGCGGATAAATACCCAAATCTAA 7198  
QY 5669 AACTCTTTTAAACGTTAAAGGACAGTATGCTGCTGTATTAAGCCCAATCAGCT 5728  
Db 7199 AACTCTTTTAAACGTTAAAGGACAGTATGCTGCTGTATTAAGCCCAATCAGCT 7258  
QY 5729 CGTAGTCTGATCTCATCAACTTCAGGAGCCACTATCTTGTGTTTAGAGAAATTTCCGGAGA 5788  
Db 7259 CGTAGTCTGATCTCATCAACTTCAGGAGCCACTATCTTGTGTTTAGAGAAATTTCCGGAGA 7318  
QY 5789 TCGCATATCGAGAAATAAGGTACGCTGATTTTAAACGTTAAAGTATTTATCTCAAGATCTCTGC 5848  
Db 7319 TCGCATATCGAGAAATAAGGTACGCTGATTTTAAACGTTAAAGTATTTATCTCAAGATCTCTGC 7378  
QY 5849 CTCGCGCGTTTCGGTGATGACGTTGAAACCTCTGACATGCGAGTCCCGGAGCGGTC 5908  
Db 7379 CTCGCGCGTTTCGGTGATGACGTTGAAACCTCTGACATGCGAGTCCCGGAGCGGTC 7438  
QY 5909 ACAGCTTGTCTGTAAGCGGATCCCGGAGCAGACAAGCCCTCAGGCGCTCAGCGGCT 5968  
Db 7439 ACAGCTTGTCTGTAAGCGGATCCCGGAGCAGACAAGCCCTCAGGCGCTCAGCGGCT 7498  
QY 5969 GTTGGCGGGTTCGCGGCGCAGCATGACCCAGTCACTAGCAGTACGCGAGTGTATACT 6028  
Db 7499 GTTGGCGGGTTCGCGGCGCAGCATGACCCAGTCACTAGCAGTACGCGAGTGTATACT 7558  
QY 6029 GGCCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAA 6088  
Db 7559 GGCCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAA 7618  
QY 6089 TACCGCAGATGCTTAAGGAGAAATACCGCATCGGCGCTCTTCCGCTTCCTCGCTCA 6148  
Db 7619 TACCGCAGATGCTTAAGGAGAAATACCGCATCGGCGCTCTTCCGCTTCCTCGCTCA 7678  
QY 6149 CTGACTCTGCGCTCGGTTCGCTGCGGCGGAGCGGTATCAGCTCACTCAAGGCGG 6208



This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 11:31:21 ; Search time 20392.8 Seconds  
(without alignments)  
17515.470 Million cell updates/sec

Title: US-10-030-390-3  
Perfect score: 8241  
Sequence: 1 agatcaacatccaagacg.....aaaataacagtattattatg 8241

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 5940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.hcg.\*
- 3: gb.in.\*
- 4: gb.ov.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.scs.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rcd.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.hgo.hum.\*
- 40: em.hgo.mus.\*
- 41: em.hgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	8241	100.0	8241	6	AX069291	AX069291 Sequence
2	7555.6	91.7	8598	6	E16036	E16036 cDNA encodi
3	7223.6	87.7	7925	12	AX178045	AX178045 Expressio
4	6726.6	81.6	8020	8	PPPIC9	Z46231 P.pastoris
5	6713.6	81.5	7757	8	PPPIC3	Z46231 P.pastoris
6	6687.6	81.2	9408	6	AX156465	AX156465 Sequence
7	6681.6	79.6	9276	6	AX149597	AX149597 Sequence
8	6561.6	75.6	9276	6	AX207941	AX207941 Sequence
9	6383.8	77.5	9776	6	AX146618	AX146618 Sequence
10	6383.8	77.5	9776	6	AX147106	AX147106 Sequence
11	6250.2	75.8	9139	12	AX178634	AX178634 Expressio
12	5709.8	69.3	10462	6	AX146621	AX146621 Sequence
13	5709.8	69.3	10462	6	AX147109	AX147109 Sequence
14	5677.8	68.9	10492	6	AX146619	AX146619 Sequence
15	5665.8	68.8	10491	6	AX147107	AX147107 Sequence
16	5657.8	68.7	10512	6	AX146620	AX146620 Sequence
17	5657.8	68.7	10512	6	AX147108	AX147108 Sequence
18	5509.6	66.9	9273	8	PPPIC9K	Z46234 P.pastoris
19	5496.6	66.7	9010	8	PPPIC3K	Z46232 P.pastoris
20	5142.8	62.4	8584	6	I63561	I63561 Sequence 66
21	5130.8	62.3	8590	6	I63563	I63563 Sequence 70
22	4565.2	55.4	8157	6	I63560	I63560 Sequence 65
C 23	2659	32.3	2659	8	ARI06448	ARI06448 Sequence
C 24	2659	32.3	2659	8	PPHIS4G	X56180 P.pastoris
C 25	2634.2	32.0	5270	12	AF027958	AF027958 Expressio
C 26	2634.2	32.0	5546	12	AF027959	AF027959 Expressio
C 27	2634.2	32.0	5565	12	AF027960	AF027960 Expressio
C 28	2634.2	32.0	5796	12	AF027961	AF027961 Expressio
C 29	2627.4	31.9	5839	12	CVU14122	U14122 Cloning vec
C 30	2621.4	31.8	5839	12	CVU14124	U1424 Cloning vec
C 31	2621.4	31.8	5284	12	CVU14123	U14123 Cloning vec
C 32	2621.4	31.8	5508	12	CVU14127	U14127 Cloning vec
C 33	2621.4	31.8	5535	12	CVU14117	U14117 Cloning vec
C 34	2621.4	31.8	5619	12	CVU14119	U14119 Cloning vec
C 35	2621.4	31.8	5821	12	CVU14118	U14118 Cloning vec
C 36	2621.4	31.8	5822	12	CVU14121	U14121 Cloning vec
C 37	2621.4	31.8	5634	12	CVU14120	U14120 Cloning vec
C 38	2621.4	31.8	5634	12	CVU14125	U14125 Cloning vec
C 39	2621.4	31.8	5936	12	CVU14116	U14116 Cloning vec
C 40	2620.2	31.8	2652	8	PPU14126	U14126 Pichia past
C 41	2325.8	28.2	3753	12	SYNTRPC	M33624 Cloning vec
C 42	2325.8	28.2	3763	12	SYNTRPA	M33622 Cloning vec
C 43	2325.8	28.2	3779	12	SYNTRPH	M32985 Expression
C 44	2324.2	28.2	3771	12	SYNTRPB	M33623 Cloning vec
C 45	2324.2	28.2	3772	12	SYNTRPD	M33625 Cloning vec

ALIGNMENTS

RESULT 1	AX069291	AX069291	8241 bp	DNA	linear	PAT 25-JAN-2001
LOCUS	Sequence 3 from Patent WO0102570.					
DEFINITION	AX069291					
ACCESSION	AX069291					
VERSION	AX069291.1	GI:12579164				
KEYWORDS						
SOURCE	synthetic construct					
ORGANISM	artificial sequences.					
REFERENCE	1					
AUTHORS	Hans,W.C., Steidler,L. and Remaut,E.R.					
TITLE	Delivery of trefol peptides					
JOURNAL	Patent: WO 0102570-A 3 11-JAN-2001;					
	Vlaams Interuniversitair Instituut voor Biotechnologie (BE)					



FEATURES		Location/Qualifiers	ORIGIN	
source	1..8241	/organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="synthetic construct : plasmid"		
Query Match	100.0%;	Score 8241; DB 6; Length 8241;		
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 8241;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;		
QY	1	AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCCATCCGACATCCACAG 60		
DB	1	AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCCATCCGACATCCACAG 60		
QY	61	GTCCATTCTCACATAGTCCCAACGCAACAGAGGGGATACATAGACAGACCGGT 120		
DB	61	GTCCATTCTCACATAGTCCCAACGCAACAGAGGGGATACATAGACAGACCGGT 120		
QY	121	TGCAAAACGAGAGCTCCACATCTCTCTCTCAACACCCACTTTTGGCCATCGAAACAC 180		
DB	121	TGCAAAACGAGAGCTCCACATCTCTCTCTCAACACCCACTTTTGGCCATCGAAACAC 180		
QY	181	AGCCCAAGTATTGGGCTTGATTTGAGTCTGCTCAATTCCTTCTATTAGGCTACTA 240		
DB	181	AGCCCAAGTATTGGGCTTGATTTGAGTCTGCTCAATTCCTTCTATTAGGCTACTA 240		
QY	241	ACACATGACTTTATTAGCT 300		
DB	241	ACACATGACTTTATTAGCT 300		
QY	301	TTTCCGAATGCAACAGCTCCGATTTACCCGCAACATCACTCCAGATGAGGCTTTCTG 360		
DB	301	TTTCCGAATGCAACAGCTCCGATTTACCCGCAACATCACTCCAGATGAGGCTTTCTG 360		
QY	361	AGTGTGGGTCAAATAGTTTCATGTTCCCAATGGCCCAAACTGACATTTAAAGCT 420		
DB	361	AGTGTGGGTCAAATAGTTTCATGTTCCCAATGGCCCAAACTGACATTTAAAGCT 420		
QY	421	GTCTTGAAACCTTAATGACAAAGCGTGATCTCATCCAAAGTGAACCTAAGTTGGTTGG 480		
DB	421	GTCTTGAAACCTTAATGACAAAGCGTGATCTCATCCAAAGTGAACCTAAGTTGGTTGG 480		
QY	481	TTGAAATGCTAACCGCCAGTTGGTCAAAAGAGAACTTCCAAAGTCCGCAATACCGTTTGT 540		
DB	481	TTGAAATGCTAACCGCCAGTTGGTCAAAAGAGAACTTCCAAAGTCCGCAATACCGTTTGT 540		
QY	541	CTTGTGTTGATTCGATGACGAATGCTCAAAATPAATCTCAATTAAGTCTTAGCGGAGTCT 600		
DB	541	CTTGTGTTGATTCGATGACGAATGCTCAAAATPAATCTCAATTAAGTCTTAGCGGAGTCT 600		
QY	601	CTCTATCGCTTCTGAACCCCGGTCACCTGTGCGGAAACCAATGCGGAAACCAACCCGCT 660		
DB	601	CTCTATCGCTTCTGAACCCCGGTCACCTGTGCGGAAACCAATGCGGAAACCAACCCGCT 660		
QY	661	TTTGGATGATATGCAATGCTCCACATGTTATGCTTCCAAAGTCTGGTGGGAATACT 720		
DB	661	TTTGGATGATATGCAATGCTCCACATGTTATGCTTCCAAAGTCTGGTGGGAATACT 720		
QY	721	GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTCTAAACCCCTTACCTTACAGCAAT 780		
DB	721	GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTCTAAACCCCTTACCTTACAGCAAT 780		
QY	781	ATATAAACAAGAGAGCTGCCCTGCTTAAACCTTTTATCATCATTTATTTAGCTT 840		
DB	781	ATATAAACAAGAGAGCTGCCCTGCTTAAACCTTTTATCATCATTTATTTAGCTT 840		
QY	841	ACTTTTCAATTTGGAGCTGGTTCCAAATGCAAGCTTTTGGATTTTAAACGCTTTTAAACGA 900		
DB	841	ACTTTTCAATTTGGAGCTGGTTCCAAATGCAAGCTTTTGGATTTTAAACGCTTTTAAACGA 900		
QY	901	CAACTTGAGAGATCAAAAACAACTAATTTATTCGAAGGATCCAAACGATGAGATTTCCT 960		

DB	901	CAACTTGAGAGATCAAAAACAACTAATTTATTCGAAGGATCCAAACGATGAGATTTCCT 960		
QY	961	TCAATTTTACTGAGATTTTTATTCGAGCATCTCCGATAGCTGCTCCAGTCAACACT 1020		
DB	961	TCAATTTTACTGAGATTTTTATTCGAGCATCTCCGATAGCTGCTCCAGTCAACACT 1020		
QY	1021	ACAACAGAGATGAAACGGCACAAATTCGGCTGAAGCTGTCACTCGGTACTCAGATTTA 1080		
DB	1021	ACAACAGAGATGAAACGGCACAAATTCGGCTGAAGCTGTCACTCGGTACTCAGATTTA 1080		
QY	1081	GAAAGGATTTGATGTTGCTGTTTCCAAACAGCACAAATAACGGTTATTG 1140		
DB	1081	GAAAGGATTTGATGTTGCTGTTTCCAAACAGCACAAATAACGGTTATTG 1140		
QY	1141	TTTATAAATACTACTATTGCGAGCATCTGCTCTAAAGAAAGAGGGGTATCTCTCGAATA 1200		
DB	1141	TTTATAAATACTACTATTGCGAGCATCTGCTCTAAAGAAAGAGGGGTATCTCTCGAATA 1200		
QY	1201	AGAGAGGCTGAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1260		
DB	1201	AGAGAGGCTGAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1260		
QY	1261	GCCCCCGGAGAGGATAAATTTGGCTTCCCGGTGTCCCGCCAGCCAGCCAGCCAGCCAG 1320		
DB	1261	GCCCCCGGAGAGGATAAATTTGGCTTCCCGGTGTCCCGCCAGCCAGCCAGCCAGCCAG 1320		
QY	1321	AGAGGTTCTCTGTTTGTGACAGTGTCCGGGATTCGGGTGCTCTCCACCCCATGGCC 1380		
DB	1321	AGAGGTTCTCTGTTTGTGACAGTGTCCGGGATTCGGGTGCTCTCCACCCCATGGCC 1380		
QY	1381	ATCGAGAACCTCAAGAAAGAAATGTCCTTCTAAGTGGGCTAGAAATTCCTTAGGG 1440		
DB	1381	ATCGAGAACCTCAAGAAAGAAATGTCCTTCTAAGTGGGCTAGAAATTCCTTAGGG 1440		
QY	1441	CGGCGCCCAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTGGGCACTTACG 1500		
DB	1441	CGGCGCCCAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTGGGCACTTACG 1500		
QY	1501	AGAAGACCGCTTCTGCTAGATTTCTAATCAAGAGATGTCAAGATGCCATTGCTCGAGAG 1560		
DB	1501	AGAAGACCGCTTCTGCTAGATTTCTAATCAAGAGATGTCAAGATGCCATTGCTCGAGAG 1560		
QY	1561	ATGAGGCTTCATTTTGTATCTTTTATTTTAACTATATATATATATATATATATATATAT 1620		
DB	1561	ATGAGGCTTCATTTTGTATCTTTTATTTTAACTATATATATATATATATATATATATAT 1620		
QY	1621	GTCAATTTTGTCTTCTGCTAGAGCTTGTCTCTGATCAGCTTATCTCGAGCTGATGAA 1680		
DB	1621	GTCAATTTTGTCTTCTGCTAGAGCTTGTCTCTGATCAGCTTATCTCGAGCTGATGAA 1680		
QY	1681	TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGAATTTTCTTGGTATTTCCC 1740		
DB	1681	TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGAATTTTCTTGGTATTTCCC 1740		
QY	1741	ACTCTCTCTCAGAGTACAGAAAGATTAGTGAGAAAGTTGCTTTGTGCAAGCTTATCGATAA 1800		
DB	1741	ACTCTCTCTCAGAGTACAGAAAGATTAGTGAGAAAGTTGCTTTGTGCAAGCTTATCGATAA 1800		
QY	1801	GCTTTAATTCGGTAGTTTATCAGATTTAAATGCTACGAGTCAAGGACCGTGTATGAA 1860		
DB	1801	GCTTTAATTCGGTAGTTTATCAGATTTAAATGCTACGAGTCAAGGACCGTGTATGAA 1860		
QY	1861	ATCTAAACAATGCGCTCATCTCCGCAACCGTCAACCTGATGCTGTAGGCATAGG 1920		
DB	1861	ATCTAAACAATGCGCTCATCTCCGCAACCGTCAACCTGATGCTGTAGGCATAGG 1920		
QY	1921	CTTGTGTATGCGGTACTGCGGGCTCTTCCGGGATATCGTCCATTCCGACAGCATCGC 1980		
DB	1921	CTTGTGTATGCGGTACTGCGGGCTCTTCCGGGATATCGTCCATTCCGACAGCATCGC 1980		
QY	1981	CAGTCACATATGCGGTGCTGTAGCGCTATATGCGTGTGATGCAATTTCTATGCGCACCGCT 2040		

Db	1981	CAGTCAC	TATGCGGTGCTGCTAGCGCTATATGCGGTTGATGCAATTTCTATGGCGACCCGT	2040
Qy	2041	TC <sup>2</sup> CGGAGCACTGTCGAGCCGCTTTGGCGCGCCCGACGTCCTGCTCGTTCCGCTACTTTGG	2100	
Db	2041	TC <sup>2</sup> CGGAGCACTGTCGAGCCGCTTTGGCGCGCCCGACGTCCTGCTCGTTCCGCTACTTTGG	2100	
Qy	2101	AGCAGCTATCGACTAGCGGATCATGCGGACACACCGGCTCTGTGGATCTATCGAATCTA	2160	
Db	2101	AGCAGCTATCGACTAGCGGATCATGCGGACACACCGGCTCTGTGGATCTATCGAATCTA	2160	
Qy	2161	AATGTAAGTTAAATCTCTAAATAATTAATAAGTCCCAAGTTTCTCCATACGAACCTTAA	2220	
Db	2161	AATGTAAGTTAAATCTCTAAATAATTAATAAGTCCCAAGTTTCTCCATACGAACCTTAA	2220	
Qy	2221	CAGANTGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTGGCCAA	2280	
Db	2221	CAGANTGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTGGCCAA	2280	
Qy	2281	TATGTTTCAGTCCCTCAGGAGTTACGTCCTGTGGAAGTGAACCTTCGGAAGGTTGAG	2340	
Db	2281	TATGTTTCAGTCCCTCAGGAGTTACGTCCTGTGGAAGTGAACCTTCGGAAGGTTGAG	2340	
Qy	2341	TGTTAACTCCGCTGTATGTAGCGGCAATATCCGTAAGTTCGGCAAGTGTGTTACCGG	2400	
Db	2341	TGTTAACTCCGCTGTATGTAGCGGCAATATCCGTAAGTTCGGCAAGTGTGTTACCGG	2400	
Qy	2401	AGGAGTAATCTCCACAACCTCTCTGGAGAGTAGGCACCAACAACACAGATCCAGCGTGT	2460	
Db	2401	AGGAGTAATCTCCACAACCTCTCTGGAGAGTAGGCACCAACAACACAGATCCAGCGTGT	2460	
Qy	2461	GTACTTGATCAACATTAAGAAAGACATTCCTCGAATTCAGAGATCAAGTCTTCAGAGCGT	2520	
Db	2461	GTACTTGATCAACATTAAGAAAGACATTCCTCGAATTCAGAGATCAAGTCTTCAGAGCGT	2520	
Qy	2521	ACTGATTTGGACATTTCCAAAGCCTGCTCGTAGGTTGCACCCGATAGGTTGTAGAGTGTG	2580	
Db	2521	ACTGATTTGGACATTTCCAAAGCCTGCTCGTAGGTTGCACCCGATAGGTTGTAGAGTGTG	2580	
Qy	2581	CAATACACTTGCCTGACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGCAT	2640	
Db	2581	CAATACACTTGCCTGACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGCAT	2640	
Qy	2641	CTTCAATTTGGCAAGCTCTGTCTGTCTATATGACAGCCACAGAACTCACTCGGAT	2700	
Db	2641	CTTCAATTTGGCAAGCTCTGTCTGTCTATATGACAGCCACAGAACTCACTCGGAT	2700	
Qy	2701	CAATACCATGTTTCAGCTTGAGACAGAAAGTCTGAGGCAACGAATCTGGATCAGCGTAT	2760	
Db	2701	CAATACCATGTTTCAGCTTGAGACAGAAAGTCTGAGGCAACGAATCTGGATCAGCGTAT	2760	
Qy	2761	TATCAGCAATAACTAGAACTTTCAGAGGCCAGCAGCATGTCAATCTACACAGGGCTG	2820	
Db	2761	TATCAGCAATAACTAGAACTTTCAGAGGCCAGCAGCATGTCAATCTACACAGGGCTG	2820	
Qy	2821	ATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTTGGACCAATA	2880	
Db	2821	ATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTTGGACCAATA	2880	
Qy	2881	TTTTGTCACTTATAGGAACAGTTTCTGTTCCGTAAGCCATAGACGCTACTGCTGGCGC	2940	
Db	2881	TTTTGTCACTTATAGGAACAGTTTCTGTTCCGTAAGCCATAGACGCTACTGCTGGCGC	2940	
Qy	2941	CTCCTGCTAGCAGATACACTTTAGACCAACCTTGTGGCAACGATAGATGACTTCTGGGG	3000	
Db	2941	CTCCTGCTAGCAGATACACTTTAGACCAACCTTGTGGCAACGATAGATGACTTCTGGGG	3000	
Qy	3001	TAAAGGTTACCATCTCTTCTTAGTGGAGATGCAAAAACAATTTCTTTGCAACCGAACCT	3060	
Db	3001	TAAAGGTTACCATCTCTTCTTAGTGGAGATGCAAAAACAATTTCTTTGCAACCGAACCT	3060	
Qy	3061	TGGCAGGAACCCAGCATCAGGAAGTGGAGAGCAGAAATGGGGTTCCACGGATAT	3120	
Db	3061	TGGCAGGAACCCAGCATCAGGAAGTGGAGAGCAGAAATGGGGTTCCACGGATAT	3120	

QY	3121	AGNGGCAACTTTCTCAATAGTCTTTGCAAAAGAGAGACAGACTACACGAGGGCAAGTCT	3180
DB	3121	AGAGGGCAACTTTCTCAATAGTCTTTGCAAAAGAGAGACAGACTACACGAGGGCAAGTCT	3180
QY	3181	CAACTTGCACAGCTCTCCGTTAGTTCATGGAATTTCTGACGTTATCTATAGAGA	3240
DB	3181	CAACTTGCACAGCTCTCCGTTAGTTCATGGAATTTCTGACGTTATCTATAGAGA	3240
QY	3241	GATCAATGGCTCTCTTAAAGGTTATCTGGCAATTCGATAGTTCTCTGGGAAAGAGCTT	3300
DB	3241	GATCAATGGCTCTCTTAAAGGTTATCTGGCAATTCGATAGTTCTCTGGGAAAGAGCTT	3300
QY	3301	CTAACACAGGTGCTTTCAAAGGACTCCATCAAACTTGCGAGTTAGTTCTTAAAGGGCTT	3360
DB	3301	CTAACACAGGTGCTTTCAAAGGACTCCATCAAACTTGCGAGTTAGTTCTTAAAGGGCTT	3360
QY	3361	TGTCACCAATTTGACGAAATTTGTCGCAATTTGTTGACTAATTCATTAATCTGTTCCG	3420
DB	3361	TGTCACCAATTTGACGAAATTTGTCGCAATTTGTTGACTAATTCATTAATCTGTTCCG	3420
QY	3421	TTTCTTGGTAGGACGACGAAGGCACTCTTCAATTTCTTGTGAGGAGGCTTTAGAAAGCT	3480
DB	3421	TTTCTTGGTAGGACGACGAAGGCACTCTTCAATTTCTTGTGAGGAGGCTTTAGAAAGCT	3480
QY	3481	CAATTTTGCACATTTCAATACGACCTTCAGAGGGACTTCTTTAGTTTGGATCTTCTT	3540
DB	3481	CAATTTTGCACATTTCAATACGACCTTCAGAGGGACTTCTTTAGTTTGGATCTTCTT	3540
QY	3541	TAGGTTGTTCTTGGTGATCTCTGGCTTGGCACTCTCCTTTCCTTCTAGTGACCTTTAGGG	3600
DB	3541	TAGGTTGTTCTTGGTGATCTCTGGCTTGGCACTCTCCTTTCCTTCTAGTGACCTTTAGGG	3600
QY	3601	ACTTCATATCCAGGTTTCTCTCAACTGTCGACGTCACACGTCACCGTACTTGGCACATCTAA	3660
DB	3601	ACTTCATATCCAGGTTTCTCTCAACTGTCGACGTCACACGTCACCGTACTTGGCACATCTAA	3660
QY	3661	CTAATGCAAAATAAATAAGTCAGCACATTTCCACAGGCTATATCTCTCTTGGATTTAGCTT	3720
DB	3661	CTAATGCAAAATAAATAAGTCAGCACATTTCCACAGGCTATATCTCTCTTGGATTTAGCTT	3720
QY	3721	CTCGAAGTTCAATCAGCTTCTCCCTTAATTTTACGGTTTCAACAAACTCTCGTCTCAATA	3780
DB	3721	CTCGAAGTTCAATCAGCTTCTCCCTTAATTTTACGGTTTCAACAAACTCTCGTCTCAATA	3780
QY	3781	ACGGTTTGGTATAAGAACCTTTCTGGAGCATTTGCTCTTACGATCCCAAGGTTGGCTTCCA	3840
DB	3781	ACGGTTTGGTATAAGAACCTTTCTGGAGCATTTGCTCTTACGATCCCAAGGTTGGCTTCCA	3840
QY	3841	TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAACCAA	3900
DB	3841	TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAACCAA	3900
QY	3901	CACCTGTTGTTTCAACACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA	3960
DB	3901	CACCTGTTGTTTCAACACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA	3960
QY	3961	GCAACTTTTTCAGTTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGACGACGAGATT	4020
DB	3961	GCAACTTTTTCAGTTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGACGACGAGATT	4020
QY	4021	GGTAGACTCCAGTTTGTCTTATAGCTTCCGGAAATAGACTTTTGGACGAGTACACCA	4080
DB	4021	GGTAGACTCCAGTTTGTCTTATAGCTTCCGGAAATAGACTTTTGGACGAGTACACCA	4080
QY	4081	GGCCCAACGAGTAATTAAGAAGTTCAGCCACCAAAAGTAGTGAATAGACCATCGGGCGGT	4140
DB	4081	GGCCCAACGAGTAATTAAGAAGTTCAGCCACCAAAAGTAGTGAATAGACCATCGGGCGGT	4140
QY	4141	CAGTAGTCAAGAGCCCAACAAATTTCACTGACAGGAACTTTTGACATCTTCAGAA	4200
DB	4141	CAGTAGTCAAGAGCCCAACAAATTTCACTGACAGGAACTTTTGACATCTTCAGAA	4200

QY 4201 GTTCGTATTTCAGTCAATTCGCGAGCATCAATAATGGGATTATACCAGAAGCAACAG 4260  
DB 4201 GTTCGTATTTCAGTCAATTCGCGAGCATCAATAATGGGATTATACCAGAAGCAACAG 4260  
QY 4261 TGGAGTGCATCTACCAACTTTGCGGTCTCAGAAAAGCAGTCTCTACTACCGC 4320  
DB 4261 TGGAGTGCATCTACCAACTTTGCGGTCTCAGAAAAGCAGTCTCTACTACCGC 4320  
QY 4321 CATTAAGTCAAACTTTCAAAATCGCCAGTGGAGAGAAAAGGCACAGCGATCTAGCAT 4380  
DB 4321 CATTAAGTCAAACTTTCAAAATCGCCAGTGGAGAGAAAAGGCACAGCGATCTAGCAT 4380  
QY 4381 TAGCGGCAAGGATGCAACTTTATCAACAGGGTCTATAGATAACCCCTAGCGCTGGGA 4440  
DB 4381 TAGCGGCAAGGATGCAACTTTATCAACAGGGTCTATAGATAACCCCTAGCGCTGGGA 4440  
QY 4441 TCATCTTTGGACAACTCTTTCTGCGCAATCTAGTCCAAATCACTTTCTATTGATACCAT 4500  
DB 4441 TCATCTTTGGACAACTCTTTCTGCGCAATCTAGTCCAAATCACTTTCTATTGATACCAT 4500  
QY 4501 TATTGTACAACTTGAGCAAGTTGTGATCAGCTCTCTCAAAATGGTCTCTGTAAACGGATG 4560  
DB 4501 TATTGTACAACTTGAGCAAGTTGTGATCAGCTCTCTCAAAATGGTCTCTGTAAACGGATG 4560  
QY 4561 ACTCAACTTGACATTAATCTGAGCTCAGTGCAGTTGAGTGAATCTGATCAGGTTGTGCA 4620  
DB 4561 ACTCAACTTGACATTAATCTGAGCTCAGTGCAGTTGAGTGAATCTGATCAGGTTGTGCA 4620  
QY 4621 GCTGCTCAGCAGCATAGGGAACACGGCTTTTCTACCAAACTCAAGGAATTAACAACCT 4680  
DB 4621 GCTGCTCAGCAGCATAGGGAACACGGCTTTTCTACCAAACTCAAGGAATTAACAACCT 4680  
QY 4681 CTGCAACACTTGCATGATCGAGTGAAGGAATGTCATCTGAAAGTCGACAGTGAG 4740  
DB 4681 CTGCAACACTTGCATGATCGAGTGAAGGAATGTCATCTGAAAGTCGACAGTGAG 4740  
QY 4741 TGAGTCTTGAAGAACTCTGAAGCGGATTTTATTAATCACTGAGTCACTCAGGAGA 4800  
DB 4741 TGAGTCTTGAAGAACTCTGAAGCGGATTTTATTAATCACTGAGTCACTCAGGAGA 4800  
QY 4801 TCCTTACCGCGGACGATCTGCGGACCTGCGAGTCCGATCACCGGCGCCACAGGTG 4860  
DB 4801 TCCTTACCGCGGACGATCTGCGGACCTGCGAGTCCGATCACCGGCGCCACAGGTG 4860  
QY 4861 CGGTGCTGGCGCTATATCCGACATCACCGATGGGAGATCGGGCTCGCCACTTCG 4920  
DB 4861 CGGTGCTGGCGCTATATCCGACATCACCGATGGGAGATCGGGCTCGCCACTTCG 4920  
QY 4921 GGCTCATCAGCGCTTTGTTTCGCGCTGGGTATGGTGGCAGGCCCCGTCGCGGGGACTGT 4980  
DB 4921 GGCTCATCAGCGCTTTGTTTCGCGCTGGGTATGGTGGCAGGCCCCGTCGCGGGGACTGT 4980  
QY 4981 TGGGGCCATCTCTTGATGCAATCTCTTGGCGGCGGCTGCTCAACGGGCTCAACC 5040  
DB 4981 TGGGGCCATCTCTTGATGCAATCTCTTGGCGGCGGCTGCTCAACGGGCTCAACC 5040  
QY 5041 TACTACTGGGCTGCTTCTTAATCAGGAGTCCGATTAAGGAGAGCGTCGAGTATCTATGA 5100  
DB 5041 TACTACTGGGCTGCTTCTTAATCAGGAGTCCGATTAAGGAGAGCGTCGAGTATCTATGA 5100  
QY 5101 TTGGAAGTATGGGAATGGTATACCCGCATCTCTCAGTGTCTTCTGAGTCTCTCTATCAGAT 5160  
DB 5101 TTGGAAGTATGGGAATGGTATACCCGCATCTCTCAGTGTCTTCTGAGTCTCTCTATCAGAT 5160  
QY 5161 TATGCCAACTAAAGCAACCGAGGAGAGATTTTCAATGGTAAATTTCTCTGACTTTTGGT 5220  
DB 5161 TATGCCAACTAAAGCAACCGAGGAGAGATTTTCAATGGTAAATTTCTCTGACTTTTGGT 5220  
QY 5221 CATCAGTATAGTCAAGTCTGAGACTATCTCGGTTATGACAGCAAGAAATGTCCTTCTTGG 5280  
DB 5221 CATCAGTATAGTCAAGTCTGAGACTATCTCGGTTATGACAGCAAGAAATGTCCTTCTTGG 5280  
QY 5281 AGACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTCTTTATCAGGAACAACTTCTTGT 5340

DB 5281 AGACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTCTTTATCAGGAACAACTTCTTGT 5340  
QY 5341 TTTCGAACTTTTTCGGTGCCTTGAACATATAAATGTAGAGTGAATATCGGGTAGGAATG 5400  
DB 5341 TTTCGAACTTTTTCGGTGCCTTGAACATATAAATGTAGAGTGAATATCGGGTAGGAATG 5400  
QY 5401 GAGCGGCAAAATGCTTTACCTTCTGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGA 5460  
DB 5401 GAGCGGCAAAATGCTTTACCTTCTGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGA 5460  
QY 5461 TGCCAACTTCAGTGAACAGTTGCTATTTCGTTCAACCATTCGAAATCCAGAGAAATCA 5520  
DB 5461 TGCCAACTTCAGTGAACAGTTGCTATTTCGTTCAACCATTCGAAATCCAGAGAAATCA 5520  
QY 5521 AAGTTGTTGTCTACTATTGATCCAAAGCAGTCGCGTCTTGAAACTGACAAATAGTGTCT 5580  
DB 5521 AAGTTGTTGTCTACTATTGATCCAAAGCAGTCGCGTCTTGAAACTGACAAATAGTGTCT 5580  
QY 5581 CGTGTGAGTGCATCTTTGATGAATAATCTAGTCTTTGATCTAAATTAATCTTGACG 5640  
DB 5581 CGTGTGAGTGCATCTTTGATGAATAATCTAGTCTTTGATCTAAATTAATCTTGACG 5640  
QY 5641 AGCCAAAGCGATAAATACCCAAATCTAAACCTTTTAAACGTTAAAGGACAAAGTATG 5700  
DB 5641 AGCCAAAGCGATAAATACCCAAATCTAAACCTTTTAAACGTTAAAGGACAAAGTATG 5700  
QY 5701 TCTGCTGTATTAAACCCCAATCAGTCTGAGTCTGATCTCATCACTTGAAGGGCAC 5760  
DB 5701 TCTGCTGTATTAAACCCCAATCAGTCTGAGTCTGATCTCATCACTTGAAGGGCAC 5760  
QY 5761 TATCTGTTTGTAGAGAAATTTCCGAGATCGATATCGAGAAAGGTACCTGATTTTA 5820  
DB 5761 TATCTGTTTGTAGAGAAATTTCCGAGATCGATATCGAGAAAGGTACCTGATTTTA 5820  
QY 5821 AAGCTGAAATTTATCTCAAGATCTCTGCTCGCGGTTTTCGGTGTATGACGGTGAACCT 5880  
DB 5821 AAGCTGAAATTTATCTCAAGATCTCTGCTCGCGGTTTTCGGTGTATGACGGTGAACCT 5880  
QY 5881 CTGACATATGAGCTCCGAGACGGTCAAGCTTGTCTGTAAAGGATCCGGGAGCAG 5940  
DB 5881 CTGACATATGAGCTCCGAGACGGTCAAGCTTGTCTGTAAAGGATCCGGGAGCAG 5940  
QY 5941 ACAAGCCGTCAGGGCGGTCAGCGGTTTGGCGGTTTTCGGGCGCAGCCATGACCCA 6000  
DB 5941 ACAAGCCGTCAGGGCGGTCAGCGGTTTGGCGGTTTTCGGGCGCAGCCATGACCCA 6000  
QY 6001 GTACGATAGCAGTACGGAGTATPACTGGCTTAACTATGCGGATCAGACGATTTGTA 6060  
DB 6001 GTACGATAGCAGTACGGAGTATPACTGGCTTAACTATGCGGATCAGACGATTTGTA 6060  
QY 6061 CTGAGAGTGCACCATATCGGTGTGAATACCGCACAGATCGTAAGGAGAAATACCGC 6120  
DB 6061 CTGAGAGTGCACCATATCGGTGTGAATACCGCACAGATCGTAAGGAGAAATACCGC 6120  
QY 6121 ATCAGCGCTCTTTCGCTTCTCTGCTCACTGCTGCTCGCTCGCTCGCTCGCTCGG 6180  
DB 6121 ATCAGCGCTCTTTCGCTTCTCTGCTCACTGCTGCTCGCTCGCTCGCTCGCTCGG 6180  
QY 6181 CGAGCGGTATCAGCTCACTCAAGCGGTATACGGTTATCCACAGATCAGGGGATAC 6240  
DB 6181 CGAGCGGTATCAGCTCACTCAAGCGGTATACGGTTATCCACAGATCAGGGGATAC 6240  
QY 6241 GCAGGAAAGAACATGTAGACAAAGCCAGCAAAAGGCCAGGAAACCGTAAAGGCCGCG 6300  
DB 6241 GCAGGAAAGAACATGTAGACAAAGCCAGCAAAAGGCCAGGAAACCGTAAAGGCCGCG 6300  
QY 6301 TTGCTGGGTTTTTCCATAGGCTCGCCCTCGCTGACGAGATCACAAATTCGAGCTCA 6360  
DB 6301 TTGCTGGGTTTTTCCATAGGCTCGCCCTCGCTGACGAGATCACAAATTCGAGCTCA 6360  
QY 6361 AGTCAGAGTGGCGAAACCCGACGAGCTATAAAGATACCAAGGCTTTTCCCTCGGAGC 6420

6361	AGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACGAGCGGTTTCCCTCGGAAGC	6421	TCCTCGTGGCGCTCTCTGTTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTC
6421	TCCTCGTGGCGCTCTCTGTTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTC	6480	TCCTCGTGGCGCTCTCTGTTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTC
6481	CCTTCCGGAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGSGTAG	6540	CCTTCCGGAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGSGTAG
6481	CCTTCCGGAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGSGTAG	6540	CCTTCCGGAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGSGTAG
6541	GTGGTTCGCTCCAAAGCTGGGCTGTGTGCAGAAACCCGCCCTTCAGCGCCGACCGCTCGGCC	6600	GTGGTTCGCTCCAAAGCTGGGCTGTGTGCAGAAACCCGCCCTTCAGCGCCGACCGCTCGGCC
6541	GTGGTTCGCTCCAAAGCTGGGCTGTGTGCAGAAACCCGCCCTTCAGCGCCGACCGCTCGGCC	6600	GTGGTTCGCTCCAAAGCTGGGCTGTGTGCAGAAACCCGCCCTTCAGCGCCGACCGCTCGGCC
6601	TTATCCGGTAACCTATCGTCTTGATGTCACACCCGGTAAGACAGACTTATCGCACCTGGCA	6660	TTATCCGGTAACCTATCGTCTTGATGTCACACCCGGTAAGACAGACTTATCGCACCTGGCA
6601	TTATCCGGTAACCTATCGTCTTGATGTCACACCCGGTAAGACAGACTTATCGCACCTGGCA	6660	TTATCCGGTAACCTATCGTCTTGATGTCACACCCGGTAAGACAGACTTATCGCACCTGGCA
6661	GCAGCCACTCGTAAACAGATTAGCAGAGCAGAGTATGTAGCGCGTGTCTCAGAGTTCCTTG	6720	GCAGCCACTCGTAAACAGATTAGCAGAGCAGAGTATGTAGCGCGTGTCTCAGAGTTCCTTG
6661	GCAGCCACTCGTAAACAGATTAGCAGAGCAGAGTATGTAGCGCGTGTCTCAGAGTTCCTTG	6720	GCAGCCACTCGTAAACAGATTAGCAGAGCAGAGTATGTAGCGCGTGTCTCAGAGTTCCTTG
6721	AAGTGTGGCCTAACTACGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTG	6780	AAGTGTGGCCTAACTACGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTG
6721	AAGTGTGGCCTAACTACGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTG	6780	AAGTGTGGCCTAACTACGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTG
6781	AAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACACCGCT	6840	AAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACACCGCT
6781	AAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACACCGCT	6840	AAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACACCGCT
6841	GGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGGCGAGAAAAAAGGATCTCAA	6900	GGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGGCGAGAAAAAAGGATCTCAA
6841	GGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGGCGAGAAAAAAGGATCTCAA	6900	GGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGGCGAGAAAAAAGGATCTCAA
6901	GAGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCAGTTAA	6960	GAGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCAGTTAA
6901	GAGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCAGTTAA	6960	GAGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCAGTTAA
6961	GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCCTTTTAAATTA	7020	GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCCTTTTAAATTA
6961	GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCCTTTTAAATTA	7020	GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCCTTTTAAATTA
7021	TGAAGTTTTAAATCAATCTAAAGTATATATAGTAAACCTTGGTCTGACAGTTACCAATGC	7080	TGAAGTTTTAAATCAATCTAAAGTATATATAGTAAACCTTGGTCTGACAGTTACCAATGC
7021	TGAAGTTTTAAATCAATCTAAAGTATATATAGTAAACCTTGGTCTGACAGTTACCAATGC	7080	TGAAGTTTTAAATCAATCTAAAGTATATATAGTAAACCTTGGTCTGACAGTTACCAATGC
7081	TTAATCAGTGAGGCACCTACTCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCTGA	7140	TTAATCAGTGAGGCACCTACTCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCTGA
7081	TTAATCAGTGAGGCACCTACTCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCTGA	7140	TTAATCAGTGAGGCACCTACTCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCTGA
7141	CTCCCCCTCGTGTAGATAACTACGATACGGAGGGCTTACCATCTGCCCCCAGTGTGCA	7200	CTCCCCCTCGTGTAGATAACTACGATACGGAGGGCTTACCATCTGCCCCCAGTGTGCA
7141	CTCCCCCTCGTGTAGATAACTACGATACGGAGGGCTTACCATCTGCCCCCAGTGTGCA	7200	CTCCCCCTCGTGTAGATAACTACGATACGGAGGGCTTACCATCTGCCCCCAGTGTGCA
7201	ATGATACCGGAGACCCACGCTCACCGCTCCAGATTATCAGCAATAAACCCAGCCAGCC	7260	ATGATACCGGAGACCCACGCTCACCGCTCCAGATTATCAGCAATAAACCCAGCCAGCC
7201	ATGATACCGGAGACCCACGCTCACCGCTCCAGATTATCAGCAATAAACCCAGCCAGCC	7260	ATGATACCGGAGACCCACGCTCACCGCTCCAGATTATCAGCAATAAACCCAGCCAGCC
7261	GGAAAGGCGGAGCGCAGAAAGTGTCTCGCAACTTTATCCGCGCTCCATCCAGTCTATTAAT	7320	GGAAAGGCGGAGCGCAGAAAGTGTCTCGCAACTTTATCCGCGCTCCATCCAGTCTATTAAT
7261	GGAAAGGCGGAGCGCAGAAAGTGTCTCGCAACTTTATCCGCGCTCCATCCAGTCTATTAAT	7320	GGAAAGGCGGAGCGCAGAAAGTGTCTCGCAACTTTATCCGCGCTCCATCCAGTCTATTAAT
7321	TGTTGCCGGGAAGCTTAGAGTAAGTAGTTTGGCCAGCTTAATAGTTTGGGAACAGTTTGGCC	7380	TGTTGCCGGGAAGCTTAGAGTAAGTAGTTTGGCCAGCTTAATAGTTTGGGAACAGTTTGGCC
7321	TGTTGCCGGGAAGCTTAGAGTAAGTAGTTTGGCCAGCTTAATAGTTTGGGAACAGTTTGGCC	7380	TGTTGCCGGGAAGCTTAGAGTAAGTAGTTTGGCCAGCTTAATAGTTTGGGAACAGTTTGGCC
7381	ATTGCTCAGGCATCTCGTGGTGTCAAGCTCGCTGTTTGGTATGGCTTCATTTCAGCTCCGGT	7440	ATTGCTCAGGCATCTCGTGGTGTCAAGCTCGCTGTTTGGTATGGCTTCATTTCAGCTCCGGT
7381	ATTGCTCAGGCATCTCGTGGTGTCAAGCTCGCTGTTTGGTATGGCTTCATTTCAGCTCCGGT	7440	ATTGCTCAGGCATCTCGTGGTGTCAAGCTCGCTGTTTGGTATGGCTTCATTTCAGCTCCGGT
7441	TCCCAACGATCAAGCGGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGTCC	7500	TCCCAACGATCAAGCGGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGTCC
7441	TCCCAACGATCAAGCGGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGTCC	7500	TCCCAACGATCAAGCGGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGTCC

	7501	TTCGGTCTCCGATCGTGTCAGAAATAAGTTGGCCGCAGTAGGTATTACATCATGAAGTTATG	7560
Qy			
Dd	7501	TTCGGTCTCCGATCGTGTCAGAAATAAGTTGGCCGCAGTAGGTATTACATCATGAAGTTATG	7560
Qy	7561	GCAGCACTCATAAATCTCTTAAGTCATGCCCATCGTAAGAATGTTTTCTGTGACTGGT	7620
Dd	7561	GCAGCACTCATAAATCTCTTAAGTCATGCCCATCGTAAGAATGTTTTCTGTGACTGGT	7620
Qy	7621	GAGTACTCAAACAAGTCATCTGAGAAATAGTGTATCGCGGACCAGATGTGCTCTTGCCC CG	7680
Dd	7621	GAGTACTCAAACAAGTCATCTGAGAAATAGTGTATCGCGGACCAGATGTGCTCTTGCCC CG	7680
Qy	7681	CGCTCAACACGGGATAATACCGCGCACATAGCAGAACATTTAAAAGTGTCTCATCTGGA	7740
Dd	7681	CGCTCAACACGGGATAATACCGCGCACATAGCAGAACATTTAAAAGTGTCTCATCTGGA	7740
Qy	7741	AACAGTTCTTCGGGGCGAAAACTCTCAAGANCTTACCGCTGTGTAGATCCAGTTCGATG	7800
Dd	7741	AACAGTTCTTCGGGGCGAAAACTCTCAAGANCTTACCGCTGTGTAGATCCAGTTCGATG	7800
Qy	7801	TAAACCACTCGTGACCCAACACTCATCTTCAGCATCTTTTACTTTTCACACGCGTTTCTGGG	7860
Dd	7801	TAAACCACTCGTGACCCAACACTCATCTTCAGCATCTTTTACTTTTCACACGCGTTTCTGGG	7860
Qy	7861	TGAGCAAAAAACAGGAAGCGAAAAATGCCGCAAAAAAGGAATAAGGGCGCACACGGAAATGT	7920
Dd	7861	TGAGCAAAAAACAGGAAGCGAAAAATGCCGCAAAAAAGGAATAAGGGCGCACACGGAAATGT	7920
Qy	7921	TGAATCACTCATCTCTCTTTTTCATATATTGAAGCATTTTATCAGGGTATTATGTCTC	7980
Dd	7921	TGAATCACTCATCTCTCTTTTTCATATATTGAAGCATTTTATCAGGGTATTATGTCTC	7980
Qy	7981	ATGAGCGGATACATATTGGAATGTATTGAAAAATAAACAATAGGGGTTCCGCGCAC A	8040
Dd	7981	ATGAGCGGATACATATTGGAATGTATTGAAAAATAAACAATAGGGGTTCCGCGCAC A	8040
Qy	8041	TTTTCCCCGAAAGTGCCACCTGACGCTCTAAGAACCATTTATTCATGACATTAAACCTAT	8100
Dd	8041	TTTTCCCCGAAAGTGCCACCTGACGCTCTAAGAACCATTTATTCATGACATTAAACCTAT	8100
Qy	8101	AAAAATAGGCGTATCAGAGCGCCTTCGTCTTCAAGAAATAATTCATGTTTGACAG C	8160
Dd	8101	AAAAATAGGCGTATCAGAGCGCCTTCGTCTTCAAGAAATAATTCATGTTTGACAG C	8160
Qy	8161	TTATCATCGATAGCTCAGCTCATGTGGTATTGGAATAGACGACATCGGGAACACT G	8220
Dd	8161	TTATCATCGATAGCTCAGCTCATGTGGTATTGGAATAGACGACATCGGGAACACT G	8220
Qy	8221	AAAAATAACAGTTATTATTCC 8241       	
Dd	8221	AAAAATAACAGTTATTATTCC 8241       	
RESULT 2			
EI6036		8598 bp DNA linear PAT 28-JUL-1999	
LOCUS			
DEFINITION		cDNA encoding improved Pichia elafin.	
ACCESSION		EI6036	
VERSION		EI6036_1 GI:5710719	
KEYWORDS		JP 1998127292-A/1.	
SOURCE		Pichia pastoris	
ORGANISM		Pichia pastoris	
REFERENCE		Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	
AUTHORS		Saccharomycetales; Saccharomycetaceae; Pichia.	
TITLE		1 (bases 1 to 8598) Tanigami,M., Yanamoto,T., Okawa,N. and Zushi.M. ELAFINS-EXPRESSION VECTOR AND PRODUCTION OF ELAFINS BY UTILIZING THE SAME	
JOURNAL		Patent: JP 1998127292-A 1 19-MAY-1998; TSUMURA & CO	
COMMENT		OS Pichia pastoris	
PN		JN JP 1998127292-A/1	

PD 19-MAY-1998  
 PF 31-OCT-1996 JP 1996304233  
 PI TANIYAMA MITSUE, YAMAMOTO TAKASHI, OKAWA NORIYUKI, PI ZUSHI  
 MAKOTO  
 PC C12N15/09, C07H21/04, C07K14/39, C12N1/19, C12P21/02, (C12N1/19, PC  
 C12R1/84),  
 PC (C12P21/02, C12R1/84);  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 FH Key Location/Qualifiers

FT source 1..8598  
 /organism='Pichia pastoris'  
 /clone='pPIC9/BLP25L'  
 FT CDS 949..1377  
 /product='improved Pichia elafin' FT  
 sig\_peptide 949..1203  
 FT mat\_peptide 1204..1374  
 /product='improved Pichia elafin' FT  
 FT misc\_feature replace(1276..1278,'atg')  
 FT /note='native Pichia elafin'

## FEATURES

source

1..8598  
 /organism='Pichia pastoris'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:4922'

## ORIGIN

Query Match 91.7%; Score 7555.6; DB 6; Length 8598;  
 Best Local Similarity 94.0%; Pred. No. 0;  
 Matches 8098; Conservative 0; Mismatches 129; Indels 385; Gaps 8;  
 QY 1 AGATCTACATCCAAAGACGAAAGGTTGAATGAACCTTTTGCCATCCGACATCCACAG 60  
 DB 1 AGATCTACATCCAAAGACGAAAGGTTGAATGAACCTTTTGCCATCCGACATCCACAG 60  
 QY 61 GTCCATTCTCACATAAGTCCCAAGCCAAACAGAGGGGATACACTAGCAGAGACCGT 120  
 DB 61 GTCCATTCTCACATAAGTCCCAAGCCAAACAGAGGGGATACACTAGCAGAGACCGT 120  
 QY 121 TGCAAAACGACGACTCCACATCTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACC 180  
 DB 121 TGCAAAACGACGACTCCACATCTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACC 180  
 QY 181 AGCCAGATTATGGGCTTGATGGAGCTCGCTCATTCCAATTCCTTATTAGGCTACTA 240  
 DB 181 AGCCAGATTATGGGCTTGATGGAGCTCGCTCATTCCAATTCCTTATTAGGCTACTA 240  
 QY 241 ACACATGACTTTATTAGCTGTCTATCTCTGCCCCCTGCGAGGTTTCATGTTGTTA 300  
 DB 241 ACACATGACTTTATTAGCTGTCTATCTCTGCCCCCTGCGAGGTTTCATGTTGTTA 300  
 QY 301 TTTCGAATGCAACAAGCTCCGCTATTACACCGGAACATCACTCCAGATGAGGGCTTTCTG 360  
 DB 301 TTTCGAATGCAACAAGCTCCGCTATTACACCGGAACATCACTCCAGATGAGGGCTTTCTG 360  
 QY 361 AGTGTGGGGTCAATAGTTTCATGTTCCCAAAATGGGCCAAACTGACAGTTTAAAGCT 420  
 DB 361 AGTGTGGGGTCAATAGTTTCATGTTCCCAAAATGGGCCAAACTGACAGTTTAAAGCT 420  
 QY 421 GTCCTGGAACTAATATGACAAAAGCGTGATCTCATCCAAGATGAACCTAAGTTTGGTTCG 480  
 DB 421 GTCCTGGAACTAATATGACAAAAGCGTGATCTCATCCAAGATGAACCTAAGTTTGGTTCG 480  
 QY 481 TTGAATGCTAACGCCAGTTGGTTCAAAAGAAACTTCGAAAAGTCCGCATACCGTTTGT 540  
 DB 481 TTGAATGCTAACGCCAGTTGGTTCAAAAGAAACTTCGAAAAGTCCGCATACCGTTTGT 540  
 QY 541 CTTGTTTGGTATTGATTCACGAATGCTCAAAAATAATCTCAATTAATGCTTAGCGAGTCT 600  
 DB 541 CTTGTTTGGTATTGATTCACGAATGCTCAAAAATAATCTCAATTAATGCTTAGCGAGTCT 600

QY 601 CTCATCGCTTCTGAAACCCCGGTGCACCTGTGCGAAACGCAAAATGGGAAACACCCGCT 660  
 DB 601 CTCATCGCTTCTGAAACCCCGGTGCACCTGTGCGAAACGCAAAATGGGAAACACCCGCT 660  
 QY 661 TTTTCGATGATATGCAATGTCTCCACATTGTATGCTTCCAGATTCTTGTGGGAATACT 720  
 DB 661 TTTTCGATGATATGCAATGTCTCCACATTGTATGCTTCCAGATTCTTGTGGGAATACT 720  
 QY 721 GCTGATAGCTAAAGCTTCAATGATCAAAATTTAACTGTCTAACCCCTACTTGACAGCAAT 780  
 DB 721 GCTGATAGCTAAAGCTTCAATGATCAAAATTTAACTGTCTAACCCCTACTTGACAGCAAT 780  
 QY 781 ATATAACAGAAAGGAGCTGCCCTGTCTTAAACCTTTTATCATCATATATAGCTT 840  
 DB 781 ATATAACAGAAAGGAGCTGCCCTGTCTTAAACCTTTTATCATCATATATAGCTT 840  
 QY 841 ACTTTTCATTAATTCGACCTGTTCCAAATTTGACAAGCTTTTGAATTTTAAAGCTTTAACA 900  
 DB 841 ACTTTTCATTAATTCGACCTGTTCCAAATTTGACAAGCTTTTGAATTTTAAAGCTTTTAAACA 900  
 QY 901 CAACTTGAGAAGATCAAAACCAACTAATTTTGAAGGATCCAAACGATGAGATTTCCT 960  
 DB 901 CAACTTGAGAAGATCAAAACCAACTAATTTTGAAGGATCCAAACGATGAGATTTCCT 960  
 QY 961 TCAATTTTACTGAGTTTATTCGACATCTCCGCAATTCCTCCGATTCAGTCCAGTCAACACT 1020  
 DB 961 TCAATTTTACTGAGTTTATTCGACATCTCCGCAATTCCTCCGATTCAGTCCAGTCAACACT 1020  
 QY 1021 ACAACAGAAAGATGAAACCGCACAAATTCGGCTGAAAGCTGTCATCGTTACTCAGATTTA 1080  
 DB 1021 ACAACAGAAAGATGAAACCGCACAAATTCGGCTGAAAGCTGTCATCGTTACTCAGATTTA 1080  
 QY 1081 GAAGGGATTTTCGATTTGCTGTTTGGCAATTTTCCACAGCACAAATAACGGGTTATTG 1140  
 DB 1081 GAAGGGATTTTCGATTTGCTGTTTGGCAATTTTCCACAGCACAAATAACGGGTTATTG 1140  
 QY 1141 TTTTAAATATCTACTATTGCCAGATTGCTCTCTAAAGAAAGGGGTATCTCTCGAGAAA 1200  
 DB 1141 TTTTAAATATCTACTATTGCCAGATTGCTCTCTAAAGAAAGGGGTATCTCTCGAGAAA 1200  
 QY 1201 AGAG---AGGCTGAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1257  
 DB 1201 AGAGCTCAAGAACCACTTAAAGGGTCCGGTTCGACCAACCGGGCTCTTGCCCGATTATC 1260  
 QY 1258 ATGGCCCCCGGAGAGGATAAATTTGGCTTCCCGGTGTCACGCCAGCCAGCCAGCCAGCCAG 1317  
 DB 1261 CTGATCCCGTGGCTTTGCTGAACCCCGCAACCGTTGCTGAAAGACACTGACTGCGCCG 1320  
 QY 1318 GAG----- 1320  
 DB 1321 GGTATCAAAAATGCTCGAAAGTTTCTGCGGTATGGCATGCTTCTGTTCCGAGTAGTGA 1380  
 QY 1321 ----- 1320  
 DB 1381 AGCTCTCGAGGCCAGCTTGGCTGCGAGTCCGAGAAAGACCCATCTCCCTAATGAGTAA 1440  
 QY 1321 ----- 1320  
 DB 1441 TATCTGATAAGAACGCTTTTATTAGCTTTTAACTACATAGACGAATAATAATACGA 1500  
 QY 1321 ----- 1320  
 DB 1501 TCGGATACATTAGCATATTAGCTCTTTTACCTCTCAATTTGCTCAAGATTAGTATCGT 1560  
 QY 1321 ----- 1320  
 DB 1561 GACCAATTTTAGATCATTTCCGTGAGAACGCCACATTTTCGCGTAATTTTAAAGTTCAA 1620  
 QY 1321 -----AGAGTTGCTGTTTGTATGACAGTGTCCGGGAT 1354  
 DB 1621 TTATTGTATGCTGACACTATATAGCTCAACTGTTTAAAGGAAGTAGTGTGATAAT 1680

QY 1355 TCCCGTGGTCTTCCACCCCATGCGCATCGAGAACACTCAAGA-----AGAAGAATGTC 1409  
Db 1681 AACGGTTCAGGTGCTTACAAATGCTTATTTCAAAATCTAGATGTTTAAACAAATTTGG 1740  
QY 1410 CTTCTAACTAGTGGGT-----AGAA 1430  
Db 1741 ATACTAACCAAGCATAGTTCTCCTGGCATCTGCCACCGTTGATATTAGTAAACTACAAGAA 1800  
QY 1431 TTCCTAGGGCGCGCGAAATTAATTCGCTTACACATGACTGTTCTCAGTTCAAGTTG 1490  
Db 1801 TTCCTAGGGCGCGCGAAATTAATTCGCTTACACATGACTGTTCTCAGTTCAAGTTG 1860  
QY 1491 GGCATCTACGAGAAGACCGGCTTCTGCTAGATTCTAATCAAGAGGATGTCAGATGCCATT 1550  
Db 1861 GGCATCTACGAGAAGACCGGCTTCTGCTAGATTCTAATCAAGAGGATGTCAGATGCCATT 1920  
QY 1551 TGCCTGAGAGATGAGGGCTTCAATTTTCACTTTTATTTTAACTATAGTATAG 1610  
Db 1921 TGCCTGAGAGATGAGGGCTTCAATTTTGAATCTTTTATTTGAACCTATATAGTATAG 1980  
QY 1611 GATTTTTTTTGTCAATTTGTTTCTTCTGTCAGAGCTTGCTCTGATCAGCCTATCTCGC 1670  
Db 1981 GATTTTTTTTGTCAATTTGTTTCTTCTGTCAGAGCTTGCTCTGATCAGCCTATCTCGC 2040  
QY 1671 AGCTGATGAATATCTTGTGTAGGGTTTGGGAAATCAATTCGAGTTTGATGTTTTCTT 1730  
Db 2041 AGCTGATGAATATCTTGTGTAGGGTTTGGGAAATCAATTCGAGTTTGATGTTTTCTT 2100  
QY 1731 GGTATTTCCCACTCTCTTTCAGAGTACAGAGATTAAGTGAGAAATGCTTGTGCAAGC 1790  
Db 2101 GGTATTTCCCACTCTCTTTCAGAGTACAGAGATTAAGTGAGAAATGCTTGTGCAAGC 2160  
QY 1791 TTATCGATPAAGCTTTAATTCGGTGTAGTTTATCACAGTTAAATTTGTAACGAGTCAGGCAC 1850  
Db 2161 TTATCGATPAAGCTTTAATTCGGTGTAGTTTATCACAGTTAAATTTGTAACGAGTCAGGCAC 2220  
QY 1851 CGTGATGAATCTAACAATCGCTTCATCGTCATCTCGGACCGTCACCTGATGCTG 1910  
Db 2221 CGTGATGAATCTAACAATCGCTTCATCGTCATCTCGGACCGTCACCTGATGCTG 2280  
QY 1911 TAGGCATAGGCTTGGTTATGCGGTACTGCGGGCTCTTGGGGGATATGTCATTCGG 1970  
Db 2281 TAGGCATAGGCTTGGTTATGCGGTACTGCGGGCTCTTGGGGGATATGTCATTCGG 2340  
QY 1971 ACAGATCGCAGTCACTATGGGTGCTGTAGCGCTATATGCTGTGATGAATTTCTAT 2030  
Db 2341 ACAGATCGCAGTCACTATGGGTGCTGTAGCGCTATATGCTGTGATGAATTTCTAT 2400  
QY 2031 GCGCACCCGTTCTCGAGCACTGTCCGACCGCTTGGCGCGCGCCAGTCCTGCTCGCTT 2090  
Db 2401 GCGCACCCGTTCTCGAGCACTGTCCGACCGCTTGGCGCGCGCCAGTCCTGCTCGCTT 2460  
QY 2091 CGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACACACACCCGTCCTGTGGATCT 2150  
Db 2461 CGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACACACACCCGTCCTGTGGATCT 2520  
QY 2151 ATCGAATCTAATGTAAGTAAATCTCTAATAATTAATTAAGTCCAGTTTCTCCATA 2210  
Db 2521 ATCGAATCTAATGTAAGTAAATCTCTAATAATTAATTAAGTCCAGTTTCTCCATA 2580  
QY 2211 CGAACCTTAAACAGCATTCGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACT 2270  
Db 2581 CGAACCTTAAACAGCATTCGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACT 2640  
QY 2271 GGTTCGGCMAATATGTTTACGTCCTCAGAGTTAGCTCTTGTGAAGTATGAATCTTCGG 2330  
Db 2641 GGTTCGGCMAATATGTTTACGTCCTCAGAGTTAGCTCTTGTGAAGTATGAATCTTCGG 2700  
QY 2331 AAGGTTGAGTGTAACTCCGCTGATTTGACGGGCATATCCGTAGCTTTGGCAAGTGTGG 2390  
Db 2701 AAGGTTGAGTGTAACTCCGCTGATTTGACGGGCATATCCGTAGCTTTGGCAAGTGTGG 2760  
QY 2391 TTGGTACCGGAGGAGTAATCTCCACAACCTCTCTGGAGAGTAGGCAACCAACACAGAT 2450

Db 2761 TTGGTACCGGAGGAGTAATCTCCACAACCTCTCTGGAGAGTAGGCAACCAACACAGAT 2820  
QY 2451 CAGCGTGTGTACTTGTATCAATTAAGAGCATCTTCGATTTGAGGATCAAGTGT 2510  
Db 2821 CAGCGTGTGTACTTGTATCAATTAAGAGAGCATCTTCGATTTGAGGATCAAGTGT 2880  
QY 2511 TCAGGAGCGCTACTGATTGGACATTTCCAAAGCCTCTCTGATGTTGCCAACCGATAGGTT 2570  
Db 2881 TCAGGAGCGCTACTGATTGGACATTTCCAAAGCCTCTCTGATGTTGCCAACCGATAGGTT 2940  
QY 2571 GTAGAGTGTCAATACACTTGGGTACATTTCAAACCTTGGCAACTGACAGCTTGGTTG 2630  
Db 2941 GTAGAGTGTCAATACACTTGGGTACATTTCAAACCTTGGCAACTGACAGCTTGGTTG 3000  
QY 2631 TGAACAGCATCTTCAATTTCTGCAAGCTCTTGTCTGATATCGACAGCCAAACAATC 2690  
Db 3001 TGAACAGCATCTTCAATTTCTGCAAGCTCTTGTCTGATATCGACAGCCAAACAATC 3060  
QY 2691 ACCTGGGAATCAATACCATGTTTCTGAGACAGAGGCTGTGAGGCAACGAAATCTGGA 2750  
Db 3061 ACCTGGGAATCAATACCATGTTTCTGAGACAGAGGCTGTGAGGCAACGAAATCTGGA 3119  
QY 2751 TCAGCGTATTTATCAGCAATTAAGTAACTTCTGAGAGCCCGCAGCGGATGTCATACTA 2810  
Db 3120 TCAGCGTATTTATCAGCAATTAAGTAACTTCTGAGAGCCCGCAGCGGATGTCATACTA 3179  
QY 2811 CACAGGCGTGTGTGTCTTGAACCATCTTCTGGCAGCAGTAAACGAATGTTTCTCT 2870  
Db 3180 CACAGGCGTGTGTGTCTTGAACCATCTTCTGGCAGCAGTAAACGAATGTTTCTCT 3239  
QY 2871 GGAACCAATATTTTGTACACTTGAACAGTTTCTGTTCCGTAAAGCATAGCAGTACT 2930  
Db 3240 GGAACCAATATTTTGTACACTTGAACAGTTTCTGTTCCGTAAAGCATAGCAGTACT 3299  
QY 2931 GCTGCGGCGCTCTCTCTAGCAGTACACTTCTAGCACCACCTTGTGGGCAACGATAGT 2990  
Db 3300 GCTGCGGCGCTCTCTCTAGCAGTACACTTCTAGCACCACCTTGTGGGCAACGATAGT 3359  
QY 2991 ACTTCTGGGGTAAGGTAACATCTTCTTGTAGTGAGATGCAAAAACAATTTCTTTGCAA 3050  
Db 3360 ACTTCTGGGGTAAGGTAACATCTTCTTGTAGTGAGATGCAAAAACAATTTCTTTGCAA 3419  
QY 3051 CAGCAACCTTGGCAGGAACCCAGCATCAGGAGTGAAGGAGAGATTCGGTTCCA 3110  
Db 3420 CAGCAACCTTGGCAGGAACCCAGCATCAGGAGTGAAGGAGAGATTCGGTTCCA 3479  
QY 3111 CAGGAATATAGAGGCCAACTTTCTCAATAGGCTTTGCAAAAACGAGAGCAGACTACCCA 3170  
Db 3480 CAGGAATATAGAGGCCAACTTTCTCAATAGGCTTTGCAAAAACGAGAGCAGACTACCCA 3539  
QY 3171 GGGCAAGTCTCACTTGGCAAGCTCTCGTGTAGTGTGAGCTTCATGGAATTTCTGAGCTTA 3230  
Db 3540 GGGCAAGTCTCACTTGGCAAGCTCTCGTGTAGTGTGAGCTTCATGGAATTTCTGAGCTTA 3599  
QY 3231 TCTATAGAGATCAATGGCTCTCTTAAAGCTTCTTGGCAATTCGCAATTCCTCTGGG 3290  
Db 3600 TCTATAGAGATCAATGGCTCTCTTAAAGCTTCTTGGCAATTCGCAATTCCTCTGGG 3659  
QY 3291 AAAGGAGCTTCAACAGAGTGTTCAAAGGCACTCCATCAAACTTTGCGAGTTAGTCT 3350  
Db 3660 AAAGGAGCTTCAACAGAGTGTTCAAAGGCACTCCATCAAACTTTGCGAGTTAGTCT 3719  
QY 3351 AAAGGAGCTTGTACCAATTTTGAACCAATTTGTGACCAATTTGGTTTGAATTCCTA 3410  
Db 3720 AAAGGAGCTTGTACCAATTTTGAACCAATTTGTGACCAATTTGGTTTGAATTCCTA 3779  
QY 3411 ATCTGTTCCGTTTCTGGATAGGACGCAAGAGGCACTTCTCAATTTCTTGTGAGGAGGCC 3470  
Db 3780 ATCTGTTCCGTTTCTGGATAGGACGCAAGAGGCACTTCTCAATTTCTTGTGAGGAGGCC 3839  
QY 3471 TTAGAACGTCATTTTGCACATTTCAATACGACCTTCAGAGGGGACTTCTTTAGTTTG 3530

Db 3840 TTAGAAACGTCAATTTTGGCAAAATCAATACGACCTTCAGAAAGGACTCTTTTAGGTTTG 3899  
Qy 3531 GATTCCTTTCTTAGGTTGTTCTTGGTGTATCCCTGGCTTGGCATCTCCTTTCTTCTTAGTG 3590  
Db 3900 GATTCCTTTCTTAGGTTGTTCTTGGTGTATCCCTGGCTTGGCATCTCCTTTCTTCTTAGTG 3959  
Qy 3591 ACCTTTAGGAGCTTCATATCAGAGTTTCTCTCACCTCTGTCGAAGCTCACACGTAAGTG 3650  
Db 3960 ACCTTTAGGAGCTTCATATCAGAGTTTCTCTCACCTCTGTCGAAGCTCACACGTAAGTG 4019  
Qy 3651 GCACATCTAATGCAAAATAAATAAGTCAGACAAATCCCAGGCTATATCTTCTCTTG 3710  
Db 4020 GCACATCTAATGCAAAATAAATAAGTCAGACAAATCCCAGGCTATATCTTCTCTTG 4079  
Qy 3711 GATTTAGCTTCTGCAAGTTCAATCAGCTTCTCCTTAATTTTAGGTTTC-AACAAACTTC 3769  
Db 4080 GATTTAGCTTCTGCAAGTTCAATCAGCTTCTCCTTAATTTTAGGTTTC-AACAAACTTC 4139  
Qy 3770 GTCGTCAAAATAACGTTTGGTATAGAACCTTCTGGAGCATGCTCTTAGCATCCACAA 3829  
Db 4140 GTCGTCAAAATAACGTTTGGTATAGAACCTTCTGGAGCATGCTCTTAGCATCCACAA 4199  
Qy 3830 GGTGGCTTCCATGGCTCTAAGACCCCTTTGATGGCCAAACAGGAAGTGGTTTCCAAAGTG 3889  
Db 4200 GGT-GCTTCCATGGCTCTAAGACCCCTTTGATGGCCAAACAGGAAGTGGTTTCCAAAGTG 4258  
Qy 3890 ACAGAAACCAACCTGTTTCAACCAACAAATTTCAAGCAGTCTCCATCACAATCCAA 3949  
Db 4259 ACAGAAACCAACCTGTTTCAACCAACAAATTTCAAGCAGTCTCCATCACAATCCAA 4318  
Qy 3950 TTCCGATACCCAGCAACTTTTGAGTTGCTCCAGATGTAGCACTTTTATACCAAAACCGTG 4009  
Db 4319 TTCCGATACCCAGCAACTTTTGAGTTGCTCCAGATGTAGCACTTTTATACCAAAACCGTG 4378  
Qy 4010 ACAGAGATTTGGTAGACTCAGTTTGTGCTTATAGCTCCGGATAGACTTTTGGTA 4069  
Db 4379 ACAGAGATTTGGTAGACTCAGTTTGTGCTTATAGCTCCGGATAGACTTTTGGTA 4438  
Qy 4070 CGAGTACACAGGCGCAAGTAGTAATAGAGAGTTCAGCCACCAAGTAGTAATAGAC 4129  
Db 4439 CGAGTACACAGGCGCAAGTAGTAATAGAGAGTTCAGCCACCAAGTAGTAATAGAC 4498  
Qy 4130 ATCCGGGCGGTGAGTCAAGAGCGCAACAAATTTCTAGCAGGGAACCTTTTGGAC 4189  
Db 4499 ATCCGGGCGGTGAGTCAAGAGCGCAACAAATTTCTAGCAGGGAACCTTTTGGAC 4558  
Qy 4190 ATCTTCAGAAAGTTCGATTCAGTAGTCAATTCGCGAGCATCAATAATGGGATTAATACC 4249  
Db 4559 ATCTTCAGAAAGTTCGATTCAGTAGTCAATTCGCGAGCATCAATAATGGGATTAATACC 4618  
Qy 4250 AGAAGCAACAGTGGAGTCAATCTACCACTTTGCGGTCTCAGAAAAAGCATAAACAGT 4309  
Db 4619 AGAAGCAACAGTGGAGTCAATCTACCACTTTGCGGTCTCAGAAAAAGCATAAACAGT 4678  
Qy 4310 TCTACTACGGCCATTTAGTGAACCTTTCAATTCGCGAGTGGAGAAAGGCAAGCAGC 4369  
Db 4679 TCTACTACGGCCATTTAGTGAACCTTTCAATTCGCGAGTGGAGAAAGGCAAGCAGC 4738  
Qy 4370 GATCTAGCATTTAGCGGCGCAAGGATGCAATTTATCAACAGGCTCCTATAGATAACCCCT 4429  
Db 4739 GATCTAGCATTTAGCGGCGCAAGGATGCAATTTATCAACAGGCTCCTATAGATAACCCCT 4798  
Qy 4430 AGCGCTGGATCATCTTTGGACAACTTTCTGCGCAATCTAGGTCCCAAAATCACTTC 4489  
Db 4799 AGCGCTGGATCATCTTTGGACAACTTTCTGCGCAATCTAGGTCCCAAAATCACTTC 4858  
Qy 4490 ATTGTATACATTTATTTGTAACCTTGCAGCAAGTTGCGATCAGCTCCCTCAAAATGGTCTTC 4549  
Db 4859 ATTGTATACATTTATTTGTAACCTTGCAGCAAGTTGCGATCAGCTCCCTCAAAATGGTCTTC 4918  
Qy 4550 TGTAACGGATGACTCAACTTGCACATTAATCTTGAAGCTCAGTCGATGAGTGAACCTTGAT 4609  
Db 4919 TGTAACGGATGACTCAACTTGCACATTAATCTTGAAGCTCAGTCGATGAGTGAACCTTGAT 4978

Qy 4610 CAGTTGTGCAAGTGGTTCAGCAGCATAGGGAACACCGCTTTTCTACCAAACTCAAGGA 4669  
Db 4979 CAGTTGTGCAAGTGGTTCAGCAGCATAGGGAACACCGCTTTTCTACCAAACTCAAGGA 5038  
Qy 4670 ATTATCAAACTCTCTCAACACATTTGGGTATGCAAGGTAGCAAGGGAATGTCACTTGAAGT 4729  
Db 5039 ATTATCAAACTCTCTCAACACATTTGGGTATGCAAGGTAGCAAGGGAATGTCACTTGAAGT 5098  
Qy 4730 CGGACAGTGTAGTGTCTTGAGAAATTTCTGAAGCGGTATTTTATCATCAGTCAGTCAG 4789  
Db 5099 CGGACAGTGTAGTGTCTTGAGAAATTTCTGAAGCGGTATTTTATCATCAGTCAGTCAG 5158  
Qy 4790 TCATCAGAGATCTCTACGCGGACGATCGTGGCCGACCTGCGAGTTCGGCATCACCGG 4849  
Db 5159 TCATCAGAGATCTCTACGCGGACGATCGTGGCCGACCTGCGAGTTCGGCATCACCGG 5206  
Qy 4850 CGCCACAGTGGGTGTCTGGCGCTATATCGCGCAATCAACCGATGGGGAAGTACGGG 4909  
Db 5207 CGCCACAGTGGGTGTCTGGCGCTATATCGCGCAATCAACCGATGGGGAAGTACGGG 5266  
Qy 4910 TCGCCACTTCGGGCTCATGAGCGCTTGTTCGCGGTGGGTATGTTGGCAGGCCCGCTGGC 4969  
Db 5267 TCGCCACTTCGGGCTCATGAGCGCTTGTTCGCGGTGGGTATGTTGGCAGGCCCGCTGGC 5326  
Qy 4970 CGGGGGAAGTGTGGGCGCCATCTCTTGCATGCACCAATTCCTTCGGGCGGCGGTCTCAA 5029  
Db 5327 CGGGGGAAGTGTGGGCGCCATCTCTTGCATGCACCAATTCCTTCGGGCGGCGGTCTCAA 5386  
Qy 5030 CGGCTCTCAACTTACTACTCGGCTGCTCTAATCGAGAGTCCGATAGGAGAGAGCGTCG 5089  
Db 5387 CGGCTCTCAACTTACTACTCGGCTGCTCTAATCGAGAGTCCGATAGGAGAGAGCGTCG 5446  
Qy 5090 AGTATCTATGATTTGGAAGTATGGGAATGGTATACCCGCACTTCTTCAGTGTCTGAAGTC 5149  
Db 5447 AGTATCTATGATTTGGAAGTATGGGAATGGTATACCCGCACTTCTTCAGTGTCTGAAGTC 5506  
Qy 5150 TCCTATCAGATTAATGCCCAACTAAAGCAACCGAGGAGAGATTTTCATGGTAAATTTCTC 5209  
Db 5507 TCCTATCAGATTAATGCCCAACTAAAGCAACCGAGGAGAGATTTTCATGGTAAATTTCTC 5566  
Qy 5210 TGACCTTTTGGTTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTATGACAGCAGAAAT 5269  
Db 5567 TGACCTTTTGGTTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTATGACAGCAGAAAT 5626  
Qy 5270 GTCTTCTTGGAGACAGTAATAAGTCCCACTAAAGAAATCCTTGTATCAGGAAAC 5329  
Db 5627 GTCTTCTTGGAGACAGTAATAAGTCCCACTAAAGAAATCCTTGTATCAGGAAAC 5686  
Qy 5330 AAACTTCTTGGTTCGAACTTTTTCGTCGCTGCTATTCGTTCAAAACCATTCGGAATC 5389  
Db 5687 AAACTTCTTGGTTCGAACTTTTTCGTCGCTGCTATTCGTTCAAAACCATTCGGAATC 5746  
Qy 5390 GGGTAGGAATGAGCGGCAAAATGCTTACCTTCGACCTTCAGAGGTATGTAGGGTTT 5449  
Db 5747 GGGTAGGAATGAGCGGCAAAATGCTTACCTTCGACCTTCAGAGGTATGTAGGGTTT 5806  
Qy 5450 GTAGATCTGATGCCAACTTCAGTCAGCAACGTTGCTATTCGTTCAAAACCATTCGGAATC 5509  
Db 5807 GTAGATCTGATGCCAACTTCAGTCAGCAACGTTGCTATTCGTTCAAAACCATTCGGAATC 5866  
Qy 5510 CAGAGAAATCAAAAGTTGTTTGTCTACTATTCGATCCAGCCAGTGGGTCTTGAACCTGAC 5569  
Db 5867 CAGAGAAATCAAAAGTTGTTTGTCTACTATTCGATCCAGCCAGTGGGTCTTGAACCTGAC 5926  
Qy 5570 AATAGTGTCTCGTGTGTTTGTAGTCACTTCTGTATGAATAATCTAGTCTTGTGATCTAAA 5629  
Db 5927 AATAGTGTCTCGTGTGTTTGTAGTCACTTCTGTATGAATAATCTAGTCTTGTGATCTAAA 5986  
Qy 5630 TAATCTTGCAGGAGCAAGGCGATAAATACCCAAATCTTAAACTCTTTTAAACCGTTAAAA 5689  
Db 5987 TAATCTTGCAGGAGCAAGGCGATAAATACCCAAATCTTAAACTCTTTTAAACCGTTAAAA 6046



QY 5690 GGACAGATATGTCTGCTGTATTAACCCCAATCAGCTCGTAGTCTGATCCTCATCAAC 5749  
Db 6047 GGACAGATATGTCTGCTGTATTAACCCCAATCAGCTCGTAGTCTCATCAAC 6106  
QY 5750 TTGAGGGACACTATCTCTTTAGAGAAATTTGGGAGATCGGATATCGAGAAAAGGTA 5809  
Db 6107 TTGAGGGACACTATCTCTTTAGAGAAATTTGGGAGATCGGATATCGAGAAAAGGTA 6166  
QY 5810 CGCTGATTTAAACGTGAAATTTATCTCAAGATCTCTGCTCGGCGGTTTCGGTGATGAC 5869  
Db 6167 CGCTGATTTAAACGTGAAATTTATCTCAAGATCTCTGCTCGGCGGTTTCGGTGATGAC 6226  
QY 5870 GGTGAARACCTTGACATGACAGTCCGAGAGCGGTACAGCTGTCTGTAGCGGAT 5929  
Db 6227 GGTGAARACCTTGACATGACAGTCCGAGAGCGGTACAGCTGTCTGTAGCGGAT 6286  
QY 5930 GCCGGAGCAGACAGCCCGTCAGCGGCTGTGCGGCTGTGCGGCTGTGCGGCGCA 5989  
Db 6287 GCCGGAGCAGACAGCCCGTCAGCGGCTGTGCGGCTGTGCGGCTGTGCGGCGCA 6346  
QY 5990 GCCATGACCCAGTCACTGAGGATAGCGGATGTATATCTGCTTAACTATGCGGCATCAG 6049  
Db 6347 GCCATGACCCAGTCACTGAGGATAGCGGATGTATATCTGCTTAACTATGCGGCATCAG 6406  
QY 6050 AGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAATACCGCACAGATGCGTAAGGA 6109  
Db 6407 AGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAATACCGCACAGATGCGTAAGGA 6466  
QY 6110 GAAATACCGCATCAGGCGCTCTTCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCG 6169  
Db 6467 GAAATACCGCATCAGGCGCTCTTCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCG 6526  
QY 6170 TTCGCTCGGCGAGCGGTATCAGTCACTCAAGGCGGTAACTAGGTATTCACAGAT 6229  
Db 6527 TTCGCTCGGCGAGCGGTATCAGTCACTCAAGGCGGTAACTAGGTATTCACAGAT 6586  
QY 6230 CAGGGATTAACGCGAGAAAGCACTGTGAGCAAAAGCGCAGCAAAAGCCAGGAACCGTA 6289  
Db 6587 CAGGGATTAACGCGAGAAAGCACTGTGAGCAAAAGCGCAGCAAAAGCCAGGAACCGTA 6646  
QY 6290 AAAAGCCGCGTGTCTGCGGTTTTTCATAGGCTCGGCGCCCTCGAGAGCATCAAAAA 6349  
Db 6647 AAAAGCCGCGTGTCTGCGGTTTTTCATAGGCTCGGCGCCCTCGAGAGCATCAAAAA 6706  
QY 6350 ATCGACGCTCAAGTCAGAGTGGGAAACCGACAGACTATAAAGATACAGGCGTTTC 6409  
Db 6707 ATCGACGCTCAAGTCAGAGTGGGAAACCGACAGACTATAAAGATACAGGCGTTTC 6766  
QY 6410 CCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCCCTCGCCTTACCGGATACCTGT 6469  
Db 6767 CCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCCCTCGCCTTACCGGATACCTGT 6826  
QY 6470 CGGCTTTCTCCCTTCGGAGAGCGGCGCTTCTCATAGCTCAGCTGAGGTATCTCA 6529  
Db 6827 CGGCTTTCTCCCTTCGGAGAGCGGCGCTTCTCATAGCTCAGCTGAGGTATCTCA 6886  
QY 6530 GTTCGGTGTAGTCTGCTTCCAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGGCCG 6589  
Db 6887 GTTCGGTGTAGTCTGCTTCCAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGGCCG 6946  
QY 6590 ACCGCTGGCCTTATCCGTTAACTATCGTCTTGTAGTCCAAACCCCGGTAAAGACAGCTTAT 6649  
Db 6947 ACCGCTGGCCTTATCCGTTAACTATCGTCTTGTAGTCCAAACCCCGGTAAAGACAGCTTAT 7006  
QY 6650 GCCCACTGGCAGCAGCACTGGTAAACAGATTAGCAGAGGAGGTATGTAGGCGGTGCTA 6709  
Db 7007 GCCCACTGGCAGCAGCACTGGTAAACAGATTAGCAGAGGAGGTATGTAGGCGGTGCTA 7066  
QY 6710 CAGAGTTCTTGAAGTGGGCTAACTAGGCTACACTAGAGGACAGATATTCGTTACT 6769  
Db 7067 CAGAGTTCTTGAAGTGGGCTAACTAGGCTACACTAGAGGACAGATATTCGTTACT 7126  
QY 6770 GCGCTCTCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGTCTTGTATCCGGCAAC 6829

Db 7127 GCGCTCTGTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGTCTTGTATCCGGAAC 7186  
QY 6830 AAACCAACCGCTGTAGCGGTGTTTTTTTGGTTGTAAGCAGCAGATTAACGCGCAGAAAA 6889  
Db 7187 AAACCAACCGCTGTAGCGGTGTTTTTTTGGTTGTAAGCAGCAGATTAACGCGCAGAAAA 7246  
QY 6890 AAGGATCTCAAGAAAGATCCTTTGATCTTTCTACCGGGTCTGACGCTCAGTGGAAACAA 6949  
Db 7247 AAGGATCTCAAGAAAGATCCTTTGATCTTTCTACCGGGTCTGACGCTCAGTGGAAACAA 7306  
QY 6950 ACTCAGTTAAAGGATTTTGGTCAGAGATTAACAAAAGGATCTTCACTAGATCCTTT 7009  
Db 7307 ACTCAGTTAAAGGATTTTGGTCAGAGATTAACAAAAGGATCTTCACTAGATCCTTT 7366  
QY 7010 TAAATTAATAAGATTTTAAATCAATTAAGATATATAGTAAAACTTGGTCTGACA 7069  
Db 7367 TAAATTAATAAGATTTTAAATCAATTAAGATATATAGTAAAACTTGGTCTGACA 7426  
QY 7070 GTTACCAATGCTTAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTCTGTTATCCA 7129  
Db 7427 GTTACCAATGCTTAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTCTGTTATCCA 7486  
QY 7130 TAGTTCCTTGACTCCCGCTGTAGATAACTACGATACGGAGGCTTACCATCTGGCC 7189  
Db 7487 TAGTTCCTTGACTCCCGCTGTAGATAACTACGATACGGAGGCTTACCATCTGGCC 7546  
QY 7190 CCAGTCTGCAATGATACCGGAGACCCACGCTACCGCTCCAGATTTATCAGCAATAA 7249  
Db 7547 CCAGTCTGCAATGATACCGGAGACCCACGCTACCGCTCCAGATTTATCAGCAATAA 7606  
QY 7250 ACCAGCCAGCGGAAGGCGGAGGAGGCTGCTTCAACTTATCCGCTCCATCC 7309  
Db 7607 ACCAGCCAGCGGAAGGCGGAGGAGGCTGCTTCAACTTATCCGCTCCATCC 7666  
QY 7310 AGTCTATTAAATGTTTCCCGGAAGCTAGAGTAAGTAGTTCCGAGTTAATAGTTCCGCA 7369  
Db 7667 AGTCTATTAAATGTTTCCCGGAAGCTAGAGTAAGTAGTTCCGAGTTAATAGTTCCGCA 7726  
QY 7370 AGTTGTTGCCATGCTGCAGGCACTGCTGTACGCTCGCTCGTTTGGTATGGCTTCAT 7429  
Db 7727 AGTTGTTGCCATGCTGCAGGCACTGCTGTACGCTCGCTCGTTTGGTATGGCTTCAT 7786  
QY 7430 TCAGTCTCCGTTCCCAACGATCAAGCGAGTTTACATGATCCCGCATGTTGTGCAAAAAAG 7489  
Db 7787 TCAGTCTCCGTTCCCAACGATCAAGCGAGTTTACATGATCCCGCATGTTGTGCAAAAAAG 7846  
QY 7490 CGGTTAGTCTCTTCCGCTCTCCGATCGTTGTGAGAAAGTTGGCCGAGTTATAC 7549  
Db 7847 CGGTTAGTCTCTTCCGCTCTCCGATCGTTGTGAGAAAGTTGGCCGAGTTATAC 7906  
QY 7550 TCATGTTATGGCAGCACTGCATAATCTCTTACTGTATGCCATCCGTAAGATGCTTT 7609  
Db 7907 TCATGTTATGGCAGCACTGCATAATCTCTTACTGTATGCCATCCGTAAGATGCTTT 7966  
QY 7610 CTGTGACTGAGTACTCAACCAAGTCAATCTGAGAAATAGTATGCGGCGACCGAGTT 7669  
Db 7967 CTGTGACTGAGTACTCAACCAAGTCAATCTGAGAAATAGTATGCGGCGACCGAGTT 8026  
QY 7670 GCTTTTCCCGCGGTCAACCGGATAATACCGGCCACATAGCAGAACTTTAAAGTGC 7729  
Db 8027 GCTTTTCCCGCGGTCAACCGGATAATACCGGCCACATAGCAGAACTTTAAAGTGC 8086  
QY 7730 TCATCAATGGAATAAGTTCTTCCGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGAT 7789  
Db 8087 TCATCAATGGAATAAGTTCTTCCGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGAT 8146  
QY 7790 CAGTTTCGATGTAAACCACTCGTGACCCCACTGTATCTTACGATCTTTTACTTTCACCA 7849  
Db 8147 CAGTTTCGATGTAAACCACTCGTGACCCCACTGTATCTTACGATCTTTTACTTTCACCA 8206  
QY 7850 GCGTTTCTGCTGAGCAAAAAAGGAAAGCAAAATGCGCAAAAGGGAATAAGGCGCA 7909



QY 1021 ACAACAGAGATGAACGACCAAAATCCGGCTGAAGCTGTATCGGTTACTCAGATTTA 1080  
DB 1012 ACTACCGAGGATGAACCTGCTCAAAATCCCTGCTGAGGCTGTATCGGTTACTCAGCTG 1071  
QY 1081 GAAGGGATTTTCGATGTTGCTGTTTTCATATTTTCCACACACACAAATAACGGGTTATTG 1140  
DB 1072 GAGGTTGACTTCGAGCTGCTGTCTGCTCAATTTCTTAATCCACCAACAGGTTTGTG 1131  
QY 1141 TTTTAAATACTACTATTATTCGACGATTCCTGCTAAAGAAAGAGGGGTATCTCTCGAGAAA 1200  
DB 1132 -----GAGGAGCTGAAGCTGAAGCTGAACCTAAATTCATCAACACT 1173  
QY 1201 AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGATCATG 1260  
DB 1174 ACTATCGCTTCTATCGCTGCTAAGGAGGAGGTTTCCCTCGAAGAAAGAGGCTGAG 1233  
QY 1261 GCCCGCCGGAGAGGATAAAATTTGCTGCTCCCGGTGTACCGCCAGCAGTGCACGGAG 1320  
DB 1234 GCTTACGTAGAGTTGA-----1249  
QY 1321 AGAGTTGCTGTTTTGATGACAGTGTCCGGGATTTCCGTGGTGTCTCCACCCCATGGCC 1380  
DB 1250 -----1249  
QY 1381 ATCGAGAACACTCAAGAGAGAAATGTCCCTTCTAATAGTGGCGTAGAATTCCTAGGG 1440  
DB 1250 -----GCTCGGG 1256  
QY 1441 CGCGCGGAATTAATTCGCTTAGACATCACTGTTCTCAGTTCAAGTTGGGCACTTACG 1500  
DB 1257 CGCGCGGAATTAATTCGCTTAGACATCACTGTTCTCAGTTCAAGTTGGGCACTTACG 1316  
QY 1501 AGAAGACCGGCTTCTAGATTCATCAAGAGGATGTCAAGTGCCTATTTGCTCGAGAG 1560  
DB 1317 AGAAGACCGGCTTCTAGATTCATCAAGAGGATGTCAAGTGCCTATTTGCTCGAGAG 1376  
QY 1561 ATGACGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1620  
DB 1377 ATGACGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1436  
QY 1621 GTCATTTTGTCTTCTCGTAGAGCTGCTCTGATCAGGCTATCTCGCAGCTGATGAA 1680  
DB 1437 GTCATTTTGTCTTCTCGTAGAGCTGCTCTGATCAGGCTATCTCGCAGCTGATGAA 1496  
QY 1681 TATCTTGTGGTAGGGTTTGGAAATCATTCGAGTTTGATGTTTCTTCTGTTATTTCCC 1740  
DB 1497 TATCTTGTGGTAGGGTTTGGAAATCATTCGAGTTTGATGTTTCTTCTGTTATTTCCC 1556  
QY 1741 ACTCCTCTTCAGAGTACAGAGATTAGTGAGAAAGTTTCTTTGTGCAAGCTTATCGATAA 1800  
DB 1557 ACTCCTCTTCAGAGTACAGAGATTAGTGAGAAAGTTTCTTTGTGCAAGCTTATCGATAA 1608  
QY 1801 GCTTTAATGCGGTAGTTTATACAGTTAAATGCTAAGCGTACAGGACCGCTGATGAA 1860  
DB 1609 GCTTTAATGCGGTAGTTTATACAGTTAAATGCTAAGCGTACAGGACCGCTGATGAA 1668  
QY 1861 ATCTAACATGCGCTCATCGTATCCTCGGACCCCTCACCCTGATGCTGATGAGATAGG 1920  
DB 1669 ATCTAACATGCGCTCATCGTATCCTCGGACCCCTCACCCTGATGCTGATGAGATAGG 1728  
QY 1921 CTTGGTTATGCGGTAGTCTCGCGGCTCTTGGGGATATGCTCCATTCGACAGCATGCG 1980  
DB 1729 CTTGGTTATGCGGTAGTCTCGCGGCTCTTGGGGATATGCTCCATTCGACAGCATGCG 1788  
QY 1981 CAGTCATATGCGGTAGTCTAGCGCTATATGCGTTGATGCAATTTCTATGCGACCCGT 2040  
DB 1789 CAGTCATATGCGGTAGTCTAGCGCTATATGCGTTGATGCAATTTCTATGCGACCCGT 1848  
QY 2041 TCTCGAGCACTGTCCGACCGCTTTGGCGCGCCCGAGTCCTGCTGCTTCTGCTACTTGG 2100  
DB 1849 TCTCGAGCACTGTCCGACCGCTTTGGCGCGCCCGAGTCCTGCTGCTTCTGCTACTTGG 1908  
QY 2101 ACCCACTATCGACTACGGCATCATGGCGACACACCCGCTCTGTGTGATCTATCGAATCTA 2160

DB 1909 AGCCACTATCGACTACGCGATCATGCGACCAACCCGCTCTGTGATCTATCGAATCTTA 1968  
QY 2161 AATGTAAGTTAAATCTTAATAATAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2220  
DB 1969 AATGTAAGTTAAATCTTAATAATAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2028  
QY 2221 CAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2280  
DB 2029 CAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2088  
QY 2281 TATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAGTGTAGTGAATCTCTGAGGTTGCGAG 2340  
DB 2089 TATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAGTGTAGTGAATCTCTGAGGTTGCGAG 2148  
QY 2341 TGTTAACTCGCTGTATTGACGGGCATATCCGTCAGTTTGGCAAGTGTGGTTGGTACCGG 2400  
DB 2149 TGTTAACTCGCTGTATTGACGGGCATATCCGTCAGTTTGGCAAGTGTGGTTGGTACCGG 2208  
QY 2401 AGAGTAATCTCCAAACTCTCTGGAGAGTAGGCACCAACACAGATCCAGCGTGT 2460  
DB 2209 AGAGTAATCTCCAAACTCTCTGGAGAGTAGGCACCAACACAGATCCAGCGTGT 2268  
QY 2461 GTACTTTGATCAACATGAAGAGCAATCTCTGATTTGCAAGGATCAAGTGTTCAGGAGCGT 2520  
DB 2269 GTACTTTGATCAACATGAAGAGCAATCTCTGATTTGCAAGGATCAAGTGTTCAGGAGCGT 2328  
QY 2521 ACTGATGAGCAATTTCAAAGCCTCTCTGATTTGCAAGGATCAAGTGTTCAGGAGCGT 2580  
DB 2329 ACTGATGAGCAATTTCAAAGCCTCTCTGATTTGCAAGGATCAAGTGTTCAGGAGCGT 2388  
QY 2581 CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGCAT 2640  
DB 2389 CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGCAT 2448  
QY 2641 CTTCAATTTTGGCAAGCTCTTGTCTGATTCGACAGCCCAAGAAATCACTCTGGAAT 2700  
DB 2449 CTTCAATTTTGGCAAGCTCTTGTCTGATTCGACAGCCCAAGAAATCACTCTGGAAT 2508  
QY 2701 CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGCAT 2760  
DB 2509 CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTTGTGAACAGCAT 2567  
QY 2761 TATCAGCAATACTAGAACTTCAGAAAGGCCAGCAGGATGTCAATCTACACAGGCGTG 2820  
DB 2568 TATCAGCAATACTAGAACTTCAGAAAGGCCAGCAGGATGTCAATCTACACAGGCGTG 2627  
QY 2821 ATGTGTCATTTGAAACCATCTCTGGCAGCAGTAAAGAACTGGTTTCTTGGACCAATA 2880  
DB 2628 ATGTGTCATTTGAAACCATCTCTGGCAGCAGTAAAGAACTGGTTTCTTGGACCAATA 2687  
QY 2881 TTTTGTACACTTAGGAACAGATTTCTTCCGTAAAGCCATAGCAGTACTGCTGGGCGC 2940  
DB 2688 TTTTGTACACTTAGGAACAGATTTCTTCCGTAAAGCCATAGCAGTACTGCTGGGCGC 2747  
QY 2941 CTCTCTAGCAGATACACTTAGCACCACCTTGTGGGCAAGCTAGATGACTTCTGGGG 3000  
DB 2748 CTCTCTAGCAGATACACTTAGCACCACCTTGTGGGCAAGCTAGATGACTTCTGGGG 2807  
QY 3001 TRAGGGTACCATCTTCTTAGGTGGAGATGCAAAACAAATTTCTTGGCAACAGCAACTT 3060  
DB 2808 TRAGGGTACCATCTTCTTAGGTGGAGATGCAAAACAAATTTCTTGGCAACAGCAACTT 2867  
QY 3061 TGGCAGGAAACCCAGCAGTCCAGGAAAGTGGAGGCAAGATTTGGGTTCCACCAAGGAATAT 3120  
DB 2868 TGGCAGGAAACCCAGCAGTCCAGGAAAGTGGAGGCAAGATTTGGGTTCCACCAAGGAATAT 2927  
QY 3121 AGAGGCAACTTCTCAATAGTCTTGGCAAGCAGAGCAGATCAACAGGCGAGTCT 3180  
DB 2928 AGAGGCAACTTCTCAATAGTCTTGGCAAGCAGAGCAGATCAACAGGCGAGTCT 2987  
QY 3181 CAACCTGCAAGCTCTCGGTTAGTGTAGCTTCATGGAATTTCTGAGCTTATCTATAGAG 3240

Db 2988 CAACCTTGAACGCTCTCGGTAGTGTAGCTTCAATGGAATTTCTGAGCTTATCTATAGAGA 3047  
 Qy 3241 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGCAATGCTCTCTGGGAAAGGAGCTT 3300  
 Db 3048 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGCAATGCTCTCTGGGAAAGGAGCTT 3107  
 Qy 3301 CTAACACAGGTGTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTT 3360  
 Db 3108 CTAACACAGGTGTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTT 3167  
 Qy 3361 TGTCAACATTTGACGACATTTGCAATTTGGTGTGATTAATTCATTAATCTGTTCCG 3420  
 Db 3168 TGTCAACATTTGACGACATTTGCAATTTGGTGTGATTAATTCATTAATCTGTTCCG 3227  
 Qy 3421 TTTTCTGTAGTAGGACGAGGAGGCTCTTCAATTTCTTGTGAGGAGGCTTTAGAAACGT 3480  
 Db 3228 TTTTCTGTAGTAGGACGAGGAGGCTCTTCAATTTCTTGTGAGGAGGCTTTAGAAACGT 3287  
 Qy 3481 CAATTTTGCACAAATCAATAGCACTTCAGAGGACTTCTTGTAGTTGGATTTCTTCT 3540  
 Db 3288 CAATTTTGCACAAATCAATAGCACTTCAGAGGAGGACTTCTTGTAGTTGGATTTCTTCT 3347  
 Qy 3541 TAGGTTGTCTCTGGTGTATCTCTGGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGG 3600  
 Db 3348 TAGGTTGTCTCTGGTGTATCTCTGGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGG 3407  
 Qy 3601 ACTTCATATCCAGGTTCTCTCCACCTGTGCTCAACAGTCACACCGTACTTGGCACAATCAA 3660  
 Db 3408 ACTTCATATCCAGGTTCTCTCCACCTGTGCTCAACAGTCACACCGTACTTGGCACAATCAA 3467  
 Qy 3661 CTAATGCAAAATAAATTAAGTCACACATCTCCAGGCTATATCTCTGCTGATTTAGCTT 3720  
 Db 3468 CTAATGCAAAATAAATTAAGTCACACATCTCCAGGCTATATCTCTGCTGATTTAGCTT 3527  
 Qy 3721 CTGCAAGTTCATCAGCT 3779  
 Db 3528 CTGCAAGTTCATCAGCT 3587  
 Qy 3780 AACCGTTTGGTATTAAGAACCTTCTGGAGCTTCTCTAGCATCCACAGGTTGGCTTCC 3839  
 Db 3588 AACCGTTTGGTATTAAGAACCTTCTGGAGCTTCTCTAGCATCCACAGGTT-GCTTCC 3846  
 Qy 3840 ATGGCTCTAAGACCTTTGATTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAACCA 3899  
 Db 3647 ATGGCTCTAAGACCTTTGATTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAACCA 3706  
 Qy 3900 ACACCTGTTTGTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCC 3959  
 Db 3707 ACACCTGTTTGTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCC 3766  
 Qy 3960 AGCACTTTTGGATTTCTCCAGATGATGACCTTTATACCAAAACCGTGACAGCAGAT 4019  
 Db 3767 AGCACTTTTGGATTTCTCCAGATGATGACCTTTATACCAAAACCGTGACAGCAGAT 3826  
 Qy 4020 TGGTAGATCCAGTTTGTCTCTATAGCTCCGGAATAGACTTTTGGAGCAGTACACC 4079  
 Db 3827 TGGTAGATCCAGTTTGTCTCTATAGCTCCGGAATAGACTTTTGGAGCAGTACACC 3886  
 Qy 4080 AGGCCCAACGATTAATAGAGAGTCCGCCAACCAAGAGTAGTAATAGACCAATCGGGCGG 4139  
 Db 3887 AGGCCCAACGATTAATAGAGAGTCCGCCAACCAAGAGTAGTAATAGACCAATCGGGCGG 3946  
 Qy 4140 TCAGTAGTCAAGACGCGCAACAAATTTCACTGACAGGAACTTTTGAATCTTTCAGAA 4199  
 Db 3947 TCAGTAGTCAAGACGCGCAACAAATTTCACTGACAGGAACTTTTGAATCTTTCAGAA 4006  
 Qy 4200 AGTTCTGATTCAGTAGTCAATTCGAGCATCAATTAATGGGATTAATACAGAGACCA 4259  
 Db 4007 AGTTCTGATTCAGTAGTCAATTCGAGCATCAATTAATGGGATTAATACAGAGACCA 4066  
 Qy 4260 GTGGAAGTCACATCTACCAACTTTGCGGTCTCAGAAAAAGCATAAACAGTTTCTACTACG 4319  
 Db 4067 GTGGAAGTCACATCTACCAACTTTGCGGTCTCAGAAAAAGCATAAACAGTTTCTACTACG 4126

Qy 4320 CCATTTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAGGCAAGCGATCTACTAGCA 4379  
 Db 4127 CCATTTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAGGCAAGCGATCTACTAGCA 4186  
 Qy 4380 TTAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTATAGATAAACCTTAGCCCTTGGG 4439  
 Db 4187 TTAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTATAGATAAACCTTAGCCCTTGGG 4246  
 Qy 4440 ATCATCTCTTTGGACAACTTTTCTGCCAAATCTTAGTCCAAATCATCTTATGATACCA 4499  
 Db 4247 ATCATCTCTTTGGACAACTTTTCTGCCAAATCTTAGTCCAAATCATCTTATGATACCA 4306  
 Qy 4500 TTATTTGACAACTTTGACGAACTTTGCGATCAGCTCTCAATTTGCTCTCTGTAACGAT 4559  
 Db 4307 TTATTTGACAACTTTGACGAACTTTGCGATCAGCTCTCAATTTGCTCTCTGTAACGAT 4366  
 Qy 4560 GACTCAACTTTGACAACTTTGAGCTCAGTTCGATTTGAGTGAACCTTATCAGCTTTGTC 4619  
 Db 4367 GACTCAACTTTGACAACTTTGAGCTCAGTTCGATTTGAGTGAACCTTATCAGCTTTGTC 4426  
 Qy 4620 AGCTGCTCAGAGCATTTAGGAAACACGCTTTTCTTACCAACTCAAGGAATTTATCAAC 4679  
 Db 4427 AGCTGCTCAGAGCATTTAGGAAACACGCTTTTCTTACCAACTCAAGGAATTTATCAAC 4486  
 Qy 4680 TCTGCAACACTTTGCTATGAGGTAGCAAGGAAATGTCTACTTGAAGTCGAGCAGTGA 4739  
 Db 4487 TCTGCAACACTTTGCTATGAGGTAGCAAGGAAATGTCTACTTGAAGTCGAGCAGTGA 4546  
 Qy 4740 GTGTAGTCTTGAGAAATTTCTGAAGCCGTTATTTTATTTATCAGTCACTCAGGAT 4799  
 Db 4547 GTGTAGTCTTGAGAAATTTCTGAAGCCGTTATTTTATTTATCAGTCACTCAGGAT 4606  
 Qy 4800 ATCTCTCAGCGGAGCATGTGCGCGACCTGCGAGGTCGGCATCACCGCGCCACAGGT 4859  
 Db 4607 ATCTCTCAGCGGAGCATGTGCGCGA-----TGCT 4639  
 Qy 4860 GCGGTGTCTGCGCTTATATCCGACATCAACCGATGGGAGATTCGGGCTCGCCACTTC 4919  
 Db 4640 GCTTGGGTGAACAGGTTTATTTCTGGCATCACTAAATATAGCATGCCG----- 4689  
 Qy 4920 GGGCTCAGAGCGCTTGTTCGGCGTGGTATGTTGGCAGCGCCCGTGGCGGGGAGCTG 4979  
 Db 4690 ----- 4689  
 Qy 4980 TTGGGCGGCATCTCCTTGATGACCACTTCTTGGCGGGGGTGTCTCAACCGGCTCAAC 5039  
 Db 4690 -----GGATTCCTTGGCGGGCGGTGCT-AAACCGGCTCAAC 4724  
 Qy 5040 CTACTACTGGGCTGTCTCTTAATCAGGAGTTCGATTAAGGAGAGCGCTCGAGTATCTATG 5099  
 Db 4725 CTAC-GCCTGGCTGTCTCTTAATCAGGAGTTCGATTAAGGAGAGCGCTCGAGTATCTATG 4783  
 Qy 5100 ATTTGGAAGTATCGGAATGCTGATACCCGATTTCTTCACTGTCTTCAAGTCTCTCTATCAGA 5159  
 Db 4784 ATTTGGAAGTATCGGAATGCTGATACCCGATTTCTTCACTGTCTTCAAGTCTCTCTATCAGA 4843  
 Qy 5160 TTATGCCCAACTTAAAGCAACCGGAGGAGGATTTTCAATGTTAAATTTCTCTGACTTTTGG 5219  
 Db 4844 TTATGCCCAACTTAAAGCAACCGGAGGAGGATTTTCAATGTTAAATTTCTCTGACTTTTGG 4903  
 Qy 5220 TCATCAGTAGTCTCGAATCTGAGACTATCTCGGTTATGACAGCAGAAATGTCTCTCTTG 5279  
 Db 4904 TCATCAGTAGTCTCGAATCTGAGACTATCTCGGTTATGACAGCAGAAATGTCTCTCTTG 4963  
 Qy 5280 GAGACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTGTGTTATCAGGAAACAACTTCTTG 5339  
 Db 4964 GAGACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTGTGTTATCAGGAAACAACTTCTTG 5023  
 Qy 5340 TTTTCGAACTTTTTCGGTCTGAACTATAAATCTAGAGTGGATATCTCGGTTAGGAAT 5399  
 Db 5024 TTTTCGAACTTTTTCGGTCTGAACTATAAATCTAGAGTGGATATCTCGGTTAGGAAT 5083

5400 GGAGCGGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTG 5459 QY  
5084 GGAGCGGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTG 5143 Db  
5460 ATGCCAACTTCAGTGAACAAGTGTCTATTTGGTTCAAAACCAATCCGAAATCCAGAGAAATC 5519 QY  
5144 ATGCCAACTTCAGTGAACAAGTGTCTATTTGGTTCAAAACCAATCCGAAATCCAGAGAAATC 5203 Db  
5520 AAAGTTGTTGTCTACTATTTGATCCAAAGCCAGTCCGGTCTTGAACCTGACAAATAGTGTGC 5579 QY  
5204 AAAGTTGTTGTCTACTATTTGATCCAAAGCCAGTCCGGTCTTGAACCTGACAAATAGTGTGC 5263 Db  
5580 TCGTGTGTTGAGGTCACTTTGTATGAATAAATCTAGTCTTTGATCTAAATATCTTTGAC 5639 QY  
5264 TCGTGTGTTGAGGTCACTTTGTATGAATAAATCTAGTCTTTGATCTAAATATCTTTGAC 5323 Db  
5640 GAGCAAGCGGATAAATAACCAAACTCTTTTAAACCTTTAAAGCAAGATAT 5699 QY  
5324 GAGCAAGCGGATAAATAACCAAACTCTTTTAAACCTTTAAAGCAAGATAT 5383 Db  
5700 GTCTGCCGTGATTAACCCCHAAATCAGTCTGACTCTGATCTCATCAATCTGAGGGGCA 5759 QY  
5384 GTCTGCCGTGATTAACCCCHAAATCAGTCTGACTCTGATCTCATCAATCTGAGGGGCA 5443 Db  
5760 CTATCTCTTTTGTAGAGAAATTTGGGGAGATCGGATATCGAGAAAAAGGTAGCTGATTTT 5819 QY  
5444 CTATCTCTTTTGTAGAGAAATTTGGGGAGATCGGATATCGAGAAAAAGGTAGCTGATTTT 5503 Db  
5820 AAAAGTGAATTTATCTCAAGATCTCTGCCCTCGCGGCTTTCGGTGATGACCGTGAACCC 5879 QY  
5504 AAAAGTGAATTTATCTCAAGATCTCTGCCCTCGCGGCTTTCGGTGATGACCGTGAACCC 5563 Db  
5880 TCTGACATCGAGCTCCCGAGACGGTCAAGCTTGTCTAAAGCGGATGCCCGGAGCA 5939 QY  
5564 TCTGACATCGAGCTCCCGAGACGGTCAAGCTTGTCTAAAGCGGATGCCCGGAGCA 5623 Db  
5940 GACAAAGCCGTCAGGCGCGTCAGCGGTGTTGGCGGTGTCGGGGCGAGCCCATGACCC 5999 QY  
5624 GACAAAGCCGTCAGGCGCGTCAGCGGTGTTGGCGGTGTCGGGGCGAGCCCATGACCC 5683 Db  
6000 AGTCAGTTCAGATAGCGAGTGTATCTGCTTAACTATCGGCATCCAGAGCAGATTGT 6059 QY  
5684 AGTCAGTTCAGATAGCGAGTGTATCTGCTTAACTATCGGCATCCAGAGCAGATTGT 5743 Db  
6060 ACTGAGAGTGACCAATATCGGTGTGAATATCCGACAGATCGGTAAAGAGAAATACCG 6119 QY  
5744 ACTGAGAGTGACCAATATCGGTGTGAATATCCGACAGATCGGTAAAGAGAAATACCG 5803 Db  
6120 CATCAGGCGCTTCCGCTTCCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGC 6179 QY  
5804 CATCAGGCGCTTCCGCTTCCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGC 5863 Db  
6180 GCGAGCGGTATCAGTCACTCAAGCGGTAAATACCGTTATCCAGAGATCAGGGGATAA 6239 QY  
5864 GCGAGCGGTATCAGTCACTCAAGCGGTAAATACCGTTATCCAGAGATCAGGGGATAA 5923 Db  
6240 CGCAGGAAAGAAATGTGAGCAAAAGGCGAGCAAAAGGCGAGAAACCGTAAAGGCGCGC 6299 QY  
5924 CGCAGGAAAGAAATGTGAGCAAAAGGCGAGCAAAAGGCGAGAAACCGTAAAGGCGCGC 5983 Db  
6300 GTTGTGCGGTTTTCATAGCTCCGCCCTCGAGAGCATCAAAAATCGACGCTC 6359 QY  
5984 GTTGTGCGGTTTTCATAGCTCCGCCCTCGAGAGCATCAAAAATCGACGCTC 6043 Db  
6360 AAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCGCCCTGGAAG 6419 QY  
6044 AAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCGCCCTGGAAG 6103 Db  
6420 CTCCTCTGTCGCTCTCTCTGTTCCGACCTCGCGCTTACCGATACCTGTCGCGCTTCT 6479 QY  
6104 CTCCTCTGTCGCTCTCTCTGTTCCGACCTCGCGCTTACCGATACCTGTCGCGCTTCT 6163 Db  
6480 CCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAAGCTGATAGCTATCTCAGTTCCGTTGTA 6539 QY

6164 CCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCCGTTGTA 6223 Db  
6540 GGTGCTTCCGCTCAAGCTGGGCTGTGTGCAAGAACCCCGCTTCAAGCGGACCGCTGGCG 6599 QY  
6224 GGTGCTTCCGCTCAAGCTGGGCTGTGTGCAAGAACCCCGCTTCAAGCGGACCGCTGGCG 6283 Db  
6600 CTTATCCGTAATCTACTGCTTGTAGTCCAAACCCGCTAAGACACAGACTTATCCGCACTGCG 6659 QY  
6284 CTTATCCGTAATCTACTGCTTGTAGTCCAAACCCGCTAAGACACAGACTTATCCGCACTGCG 6343 Db  
6660 AGCAGCCACTGTGTAAACAGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTT 6719 QY  
6344 AGCAGCCACTGTGTAAACAGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTT 6403 Db  
6720 GAAGTGGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGCTATCTCGCTCTGCT 6779 QY  
6404 GAAGTGGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGCTATCTCGCTCTGCT 6463 Db  
6780 GAAGCAGATTACTTTCGGAAAAAGAGTTGTTAGTCTTGTATCCGGCAAAACCAACCCGCG 6839 QY  
6464 GAAGCAGATTACTTTCGGAAAAAGAGTTGTTAGTCTTGTATCCGGCAAAACCAACCCGCG 6523 Db  
6840 TGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGAGATTACCGCAGAAAAAGGATCTCA 6899 QY  
6524 TGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGAGATTACCGCAGAAAAAGGATCTCA 6583 Db  
6900 AGAAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAGAAACCTCACTGTTA 6959 QY  
6584 AGAAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAGAAACCTCACTGTTA 6643 Db  
6960 AGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTA 7019 QY  
6644 AGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTA 6703 Db  
7020 ATCAAGTTTTTAAATCAATCTAAAGTATATATAGTAAATCTTGGTCTGACAGTTACCAATG 7079 QY  
6704 ATCAAGTTTTTAAATCAATCTAAAGTATATATAGTAAATCTTGGTCTGACAGTTACCAATG 6763 Db  
7080 CTTTAAATCAGTAGGCACTATCTCAGCGATCTGTCTATTTGGTTGATCTCAGTAGTTCCCTG 7139 QY  
6764 CTTTAAATCAGTAGGCACTATCTCAGCGATCTGTCTATTTGGTTGATCTCAGTAGTTCCCTG 6823 Db  
7140 ACTCCCGCTGCTAGTAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCTG 7199 QY  
6824 ACTCCCGCTGCTAGTAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCTG 6883 Db  
7200 AATGATACCGCGAGACCCAGCTCACCGGCTCCAGATTATCAGCAATAAAACCGAGCAGC 7259 QY  
6884 AATGATACCGCGAGACCCAGCTCACCGGCTCCAGATTATCAGCAATAAAACCGAGCAGC 6943 Db  
7260 CGGAAGGCGCGAGCGAGAGTGGTCTCGCACTTTATCGGCTCCATCCAGTCTATTAA 7319 QY  
6944 CGGAAGGCGCGAGCGAGAGTGGTCTCGCACTTTATCGGCTCCATCCAGTCTATTAA 7003 Db  
7320 TTGTTGCGGGAAGCTAGATAGTTAGTTGCGCAGTTAATAGTTTGGCAACCGTTGTC 7379 QY  
7004 TTGTTGCGGGAAGCTAGATAGTTAGTTGCGCAGTTAATAGTTTGGCAACCGTTGTC 7063 Db  
7380 CATTTGTCAGGCAATCGTGGTCAAGCTGCTGTTTGGTATGGCTTCATTCAGTCCCG 7439 QY  
7064 CATTTGTCAGGCAATCGTGGTCAAGCTGCTGCTGTTTGGTATGGCTTCATTCAGTCCCG 7123 Db  
7440 TTCCCAACGATCAAGCGAGTTACATGATCCCGCATTTGTGCAAAAAGCGTTAGCTC 7499 QY  
7124 TTCCCAACGATCAAGCGAGTTACATGATCCCGCATTTGTGCAAAAAGCGTTAGCTC 7183 Db  
7500 CTTCCGTCCTCCGATCGTTCTCAGAAAGTAAAGTTGGCGCAGTGTATCTCATGTTAT 7559 QY  
7184 CTTCCGTCCTCCGATCGTTCTCAGAAAGTAAAGTTGGCGCAGTGTATCTCATGTTAT 7243 Db  
7560 GCGAGCACTGCAATTTCTTACTGTCTCATGCCATCGTAAAGTGTCTTCTGTCGCTG 7619 QY

Db 7244 GGCAGCATGCTAATCTCTTCTACTGTGTCATGCCATCGTAAGATGCTTTTCTGTGACTGG 7303  
Qy 7620 TGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTATCGCGGACCGAGTTGCTCTTTCGCC 7679  
Db 7304 TGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTATCGCGGACCGAGTTGCTCTTTCGCC 7363  
Qy 7680 GCGGTCAACACGGGATATACCGCGGCACATAGCAGAACTTTAAAGTGTCTCATCTATGG 7739  
Db 7364 GCGGTCAACACGGGATATACCGCGGCACATAGCAGAACTTTAAAGTGTCTCATCTATGG 7423  
Qy 7740 AAAAGTTCTTCGGGGGAAAACCTCTCAAGGATCTTACCGTGTGTAGATCCAGTTTCGAT 7799  
Db 7424 AAAAGTTCTTCGGGGGAAAACCTCTCAAGGATCTTACCGTGTGTAGATCCAGTTTCGAT 7483  
Qy 7800 GTAAACCACTCGTGCACCCACTGATCTTTCAGCATCTTTTACTTTTCCACGAGCTTTCTGG 7859  
Db 7484 GTAAACCACTCGTGCACCCACTGATCTTTCAGCATCTTTTACTTTTCCACGAGCTTTCTGG 7543  
Qy 7860 GTGACCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAGGGCGGACACGGAAATG 7919  
Db 7544 GTGACCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAGGGCGGACACGGAAATG 7603  
Qy 7920 TTGAATCTCATCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATTTGCT 7979  
Db 7604 TTGAATCTCATCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATTTGCT 7663  
Qy 7980 CATGAGCGGATACATATTGTAATGTTTGAATAAATAAAGGGAATAGGGGTTCCGCGCAC 8039  
Db 7664 CATGAGCGGATACATATTGTAATGTTTGAATAAATAAAGGGAATAGGGGTTCCGCGCAC 7723  
Qy 8040 ATTTCGCCGAAAGTGCCACTGACGCTTAAGAAACCATTTATATCATGACATTAACCTA 8099  
Db 7724 ATTTCGCCGAAAGTGCCACTGACGCTTAAGAAACCATTTATATCATGACATTAACCTA 7783  
Qy 8100 TAAAAATAGGCGTATCAGAGGCGCTTTGCTCTCAAGAAATATTTCTCATGTTTGACAG 8159  
Db 7784 TAAAAATAGGCGTATCAGAGGCGCTTTGCTCTCAAGAAATATTTCTCATGTTTGACAG 7843  
Qy 8160 CTTATCATGTAAGTCACTCATGTTGGTATGTAATAGAGCGGACGAGATCGGGAACACT 8219  
Db 7844 CTTATCATGTAAGTCACTCATGTTGGTATGTAATAGAGCGGACGAGATCGGGAACACT 7903  
Qy 8220 GAAAAATAACAGTTATTTATTCG 8241  
Db 7904 GAAAAATAACAGTTATTTATTCG 7925

## RESULT 4

LOCUS PPPIC9 8020 bp DNA linear PLN 22-OCT-1994  
DEFINITION P.pastoris DNA for pPIC9 expression vector.  
ACCESSION Z46233  
VERSION Z46233.1 GI:559514  
KEYWORDS beta-lactamase.  
SOURCE Pichia pastoris  
ORGANISM Pichia pastoris  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Pichia.  
REFERENCE 1 (bases 1 to 7957)  
AUTHORS Scorer,C.A., Clare,J.J., McCombie,W.R., Romanos,M.A. and Sreekrishna,K.  
TITLE Rapid selection using G418 of high copy number transformants of Pichia pastoris for high-level foreign gene expression  
JOURNAL Biotechnology (N.Y.) 12 (2), 181-184 (1994)  
MEDLINE 94153511  
PUBMED 7764433  
REFERENCE 2 (bases 1 to 8020)  
AUTHORS Romanos,M.A.  
TITLE Direct Substitution  
JOURNAL Submitted (18-OCT-1994) Romanos M. A., Wellcome Research Laboratories, Biotechnology Division, Langley Court, Beckenham Kent UK

source  
1..8020  
/organism="Pichia pastoris"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4922"  
/clone="pPIC3"  
misc\_feature  
1..8020  
/standard\_name="pPIC9"  
/function="Pichia alpha-factor secretion vector"  
/label=pPIC9  
terminator  
1..354  
/standard\_name="AOX1"  
complement (1250..3280)  
/standard\_name="HIS4"  
CDS  
1..354  
/codon\_start=1  
/protein\_id="CAA86304.1"  
/db\_xref="GI:559515"  
/db\_xref="GOA:Q12201"  
/db\_xref="SPTREMBL:Q12201"  
translation="MTFLLPAYASVAEFDNSLSLVGKAVFPYAADQHNLIKFTQST  
ELQNVQVSSVTSDQFELIDNLLKLYNNGINIEVILDLDLAERVQVMI PGARVIYR  
TLIDKVASLPANASIAVDFSPPLGDLKSFVTSRTVAFSETAKLVDTSTVASGII  
PLIDARQLTTEVELSEDVKKFVSEILLASLTTRDPGLFTTLVADSSNYSGLVYS  
KKSIPEARITQGYQSRHGLWYKGTSGTOKLGLIELDCDGDCLKFVVEONGVGE  
CHERTSCFGSGKGLRANAEFCGIVRAMLKVLIPNGVLTTFCLNKKIREADELAE  
AKSDIAWECADLFYFALVRCAYKIVTLDVERNLDKMSLVTRKGDADKPGYTKQ  
PKESKPEVPSSEGRILEKIDVSKASSEIEDALRRPIQKTEIMLVKPIVDNVQ  
NGDKALLBELTAFDFGVALKTVLBAFPEELMQLPDNVKRAIDLSIDNVRKHEAQT  
PLQVETCFGVVCSRFARPIEKVGLYIPGGTALPSTLMGLVPKAVAGCKEIVFASP  
ETKQDTPLEVIYVAHVKGAKCIVLAGGAQAAVAMAYGTETVPKCDKI FQGNQFVTA  
AKMVQNTSALCSIDMPAGPSEVIVTADKVDPDFVASDLLLLKLVNVLPR"  
3589..4036  
/standard\_name="3' AOX1"  
complement (5606..6466)  
/product="beta-lactamase"  
misc\_feature  
6766..7703  
/standard\_name="Ampr"  
sig\_peptide  
7727..7979  
/standard\_name="AOX1"  
/standard\_name="alpha factor leader"

## ORIGIN

Query Match 81.6%; Score 6726.6; DB 8; Length 8020;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 6784; Conservative 0; Mismatches 4; Indels 15; Gaps 4;  
Qy 1440 GCGGCGCGAATTAATTCGCTTACAGATGACTGTTCCTCAGTTCCGGCACTTAC 1499  
Db 1 GCGGCGCGAATTAATTCGCTTACAGATGACTGTTCCTCAGTTCCGGCACTTAC 60  
Qy 1500 GAGAGACCGGTCTGTCTAGATTTCTAATCAAGAGATGTCAGATGCCATTTGCTGAGA 1559  
Db 61 GAGAGACCGGTCTGTCTAGATTTCTAATCAAGAGATGTCAGATGCCATTTGCTGAGA 120  
Qy 1560 GATGAGCGCTTCATTTTGTATCTTTTATTTATTTGTAACCTATATAGTATAGATTTT 1619  
Db 121 GATGAGCGCTTCATTTTGTATCTTTTATTTATTTGTAACCTATATAGTATAGATTTT 180  
Qy 1620 TGTCAATTTTGTCTTCTCGTACGAGCTGTCTGATCAGCCTATCTCGCAGCTGAGA 1679  
Db 181 TGTCAATTTTGTCTTCTCGTACGAGCTGTCTGATCAGCCTATCTCGCAGCTGAGA 240  
Qy 1680 ATATCTTGTGTAGGGGTTTGGGAAATCAATTCAGATTTGATGTTTTTCTTGGTATTTC 1739  
Db 241 ATATCTTGTGTAGGGGTTTGGGAAATCAATTCAGATTTGATGTTTTTCTTGGTATTTC 300  
Qy 1740 CACTCTCTTTCAGAGTACAGAAATTAAGTGAAGTTTGGTTTGCAGCTTATCGATA 1799  
Db 301 CACTCTCTTTCAGAGTACAGAAATTAAGTGAAGTTTGGTTTGCAGCTTATCGATA 360  
Qy 1800 AGCTTTAATCGGTAGTTTATCACAGTTAAATTCGTAACGAGTCAGCACCCTGTATGA 1859

Qy	1860	AATCTAACAAATGGCGCTCATCGTCACTCTCGGCAACCGTACACCTGGAGTGTGTAGGCATAG	1919
Db	421	AACTCTAACAAATGGCGCTCATCGTCACTCTCGGCAACCGTACACCTGGAGTGTGTAGGCATAG	480
Qy	1920	GCCTTGGTTATGCGGTTACTGCGGGCTCTTTGGGGATATCGTCCTATTCGACAGCATCG	1979
Db	481	GCCTTGGTTATGCGGTTACTGCGGGCTCTTTGGGGATATCGTCCTATTCGACAGCATCG	540
Qy	1980	CCAGTCACTAATGCGGTGCTGCTAGCGCTATATAGCGTTGATGCAATTTCTATGCGCACCGG	2039
Db	541	CCAGTCACTAATGCGGTGCTGCTAGCGCTATATAGCGTTGATGCAATTTCTATGCGCACCGG	600
Qy	2040	TTCTCGGAGCAGCTGTCGACCGCTTTTGGCGCGCCGACGTCGTCCTGCTACTTGG	2099
Db	601	TTCTCGGAGCAGCTGTCGACCGCTTTTGGCGCGCCGACGTCGTCCTGCTACTTGG	660
Qy	2100	GAGCCACTATCGACTAGCGGATCATGCGGACCAACCCGCTCTGTGGATCTATCGAATCT	2159
Db	661	GAGCCACTATCGACTAGCGGATCATGCGGACCAACCCGCTCTGTGGATCTATCGAATCT	720
Qy	2160	AAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATACGAACTT	2219
Db	721	AAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATACGAACTT	780
Qy	2220	ACAGCAATTCGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCTGTGTGGCCA	2279
Db	781	ACAGCAATTCGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCTGTGTGGCCA	840
Qy	2280	ATATGTTTCAGTCCCTCAGGAGTTACGTCCTGTGAAGTGAATGAATCTTGGAAAGTTGCA	2339
Db	841	ATATGTTTCAGTCCCTCAGGAGTTACGTCCTGTGAAGTGAATGAATCTTGGAAAGTTGCA	900
Qy	2340	GTGTTAACTCCGCTGTATTCAGCGGCATATCCGTTAGCTTGGCAAAGTGTGTTGGTACCG	2399
Db	901	GTGTTAACTCCGCTGTATTCAGCGGCATATCCGTTAGCTTGGCAAAGTGTGTTGGTACCG	960
Qy	2400	GAGGAGTAATCTCCACAACTCTCTGAGAGTAGGCAACCAACCAACAGATCCAGCGTGT	2459
Db	961	GAGGAGTAATCTCCACAACTCTCTGAGAGTAGGCAACCAACCAACAGATCCAGCGTGT	1020
Qy	2460	TGTACTTGATCAACATPAAGAGAGCATCTCCGATTTGCAAGTCAAGTCTTCAGAGCG	2519
Db	1021	TGTACTTGATCAACATPAAGAGAGCATCTCCGATTTGCAAGTCAAGTCTTCAGAGCG	1080
Qy	2520	TACTGATGACATTTCCAAAGCGCTGCTCGTAGTTTGCAACCGATAGGTTGTAGAGTGT	2579
Db	1081	TACTGATGACATTTCCAAAGCGCTGCTCGTAGTTTGCAACCGATAGGTTGTAGAGTGT	1140
Qy	2580	GCAATACACTTGGTACAAATTTCAACCCTTGGCAACGCAACGCTTGGTTGTGAAACAGCA	2639
Db	1141	GCAATACACTTGGTACAAATTTCAACCCTTGGCAACGCAACGCTTGGTTGTGAAACAGCA	1200
Qy	2640	TCCTCAATTTGCGCAAGCTCCTTGTCTGTATATCGACAGCCCAACAGAAATCACTCGGAA	2699
Db	1201	TCCTCAATTTGCGCAAGCTCCTTGTCTGTATATCGACAGCCCAACAGAAATCACTCGGAA	1260
Qy	2700	TCAATACACTTTCAGCTTGAGCAGAAAGTCTGAGGCAACGAAATCTGATCAGCGTAT	2759
Db	1261	TCAATACACTTTCAGCTTGAGCAGAAAGTCTGAGGCAACGAAATCTGATCAGCGTAT	1319
Qy	2760	TTATCAGCAATAACTAGAACTTTCAGAAAGTCTGAGGCAACGAAATCTGATCAGCGTAT	2819
Db	1320	TTATCAGCAATAACTAGAACTTTCAGAAAGTCTGAGGCAACGAAATCTGATCAGCGTAT	1379
Qy	2820	GATGTGTCAATTTGAAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTGACACCAAT	2879
Db	1380	GATGTGTCAATTTGAAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTGACACCAAT	1439
Qy	2880	ATTTTGTGCACATTTAGGACAGTTTCTGTTCCGTAAGCCATAGCAGTACTGCTCGGCG	2939
Db	1440	ATTTTGTGCACATTTAGGACAGTTTCTGTTCCGTAAGCCATAGCAGTACTGCTCGGCG	1499
Qy	2940	CCTCCTGCTAGCAGATACACTTTAGCAACCAACCTTGTGGCAACGATAGACTCTTGGG	2999

Db	1500	CCCTCCCTGCTAGCAGCATACACTTAGCACCAACCTTGTGGCAACGATAGATGACTTCTGGG	1559
Qy	3000	GTAAGGAGTACCATCTCTTTAGTGTGAGATGCAAAACAAATTTCTTTTGCACACGAACT	3059
Db	1560	GTAAGGATACCATCTCTTTAGTGTGAGATGCAAAACAAATTTCTTTTGCACACGAACT	1619
Qy	3060	TTGCGAGGAACACCCAGCATCAGGGAAGTGGGAAGGAGAGAAATTCGGGTTCCACAGGAATA	3119
Db	1620	TTGCGAGGAACACCCAGCATCAGGGAAGTGGGAAGGAGAGAAATTCGGGTTCCACAGGAATA	1679
Qy	3120	TAGAGGCCAACTTTCTCAATAGTCTTTGCAAAACGAGACGAGACTACACAGGGCGAAGTC	3179
Db	1680	TAGAGGCCAACTTTCTCAATAGTCTTTGCAAAACGAGACGAGACTACACAGGGCGAAGTC	1739
Qy	3180	TCAACTTTGCAACGCTCCGGTTAGTGTAGCTTCAATGGAATTTCTCTGACGTTATCTATAGAG	3239
Db	1740	TCAACTTTGCAACGCTCCGGTTAGTGTAGCTTCAATGGAATTTCTCTGACGTTATCTATAGAG	1799
Qy	3240	AGATCAATGCTCTCTTTAAGCTTATCTGGCAATTTGCATAAGTTCCTCTGGGAAGGAGCT	3299
Db	1800	AGATCAATGCTCTCTTTAAGCTTATCTGGCAATTTGCATAAGTTCCTCTGGGAAGGAGCT	1859
Qy	3300	TCTAACAACAGGTGTCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTATAAAGGGCT	3359
Db	1860	TCTAACAACAGGTGTCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTATAAAGGGCT	1919
Qy	3360	TTGTCACCAATTTTGACGAACATTTGTCGACAAATTTGGTTTGACTAATTTCCATAATCTCTTCC	3419
Db	1920	TTGTCACCAATTTTGACGAACATTTGTCGACAAATTTGGTTTGACTAATTTCCATAATCTCTTCC	1979
Qy	3420	GTTTTCCTGGATAGGACGACGAAGGGCATCTTCAATTTCTTGTGAGGAGGCCCTTAGAAACG	3479
Db	1980	GTTTTCCTGGATAGGACGACGAAGGGCATCTTCAATTTCTTGTGAGGAGGCCCTTAGAAACG	2039
Qy	3480	TCAATTTTGCACAAATTCATACAGACCTTCAGAAGGAGCATCTTCTTTAGTTTGGATCTTCT	3539
Db	2040	TCAATTTTGCACAAATTCATACAGACCTTCAGAAGGAGCATCTTCTTTAGTTTGGATCTTCT	2099
Qy	3540	TTAGTTTGTTCCTTTGGTGTATCTGGCTTGGCATCTCTTTCCTCTCTAGTGACCTTTAGG	3599
Db	2100	TTAGTTTGTTCCTTTGGTGTATCTGGCTTGGCATCTCTTTCCTCTCTAGTGACCTTTAGG	2159
Qy	3600	GACTTCATATCCAGGTTTCTCTCCACTCGTCCAAAGTCAACCGTACTTTGGGCACATCTA	3659
Db	2160	GACTTCATATCCAGGTTTCTCTCCACTCGTCCAAAGTCAACCGTACTTTGGGCACATCTA	2219
Qy	3660	ACTAATGCAAAATAAATAAGTCAGACATTCCTCCAGGCTATATCTTCTTTGGATTTAGCT	3719
Db	2220	ACTAATGCAAAATAAATAAGTCAGACATTCCTCCAGGCTATATCTTCTTTGGATTTAGCT	2279
Qy	3720	TCTGCAAGTTCAATCAGCTTCTCCCTAAATTTTAGCGTTC-AACAAAACCTTCGTCGTCAAA	3778
Db	2280	TCTGCAAGTTCAATCAGCTTCTCCCTAAATTTTAGCGTTC-AACAAAACCTTCGTCGTCAAA	2339
Qy	3779	TAAACCGTTTGGTATAGAAACCTTCTGAGGACATTTGCTCTTACGATCCCAAGGTGCGTTC	3838
Db	2340	TAAACCGTTTGGTATAGAAACCTTCTGAGGACATTTGCTCTTACGATCCCAAGGT-GGTT	2398
Qy	3839	CATGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAAGTCCGTTTCAAGTGCAGAAACCC	3898
Db	2399	CATGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAAGTCCGTTTCAAGTGCAGAAACCC	2458
Qy	3899	AACACCTGTTTGTTCACACAAATTTCAAGAGTGTCTCCATCAATCCAAATTCGATACC	3958
Db	2459	AACACCTGTTTGTTCACACCAAAATTTCAAGAGTGTCTCCATCAATTCGATACC	2518
Qy	3959	CAGCAACTTTTGATTTGCTCCAGATGTAGCACCTTTATACCAAAACCGTGACGACGAGA	4018
Db	2519	CAGCAACTTTTGATTTGCTCCAGATGTAGCACCTTTATACCAAAACCGTGACGACGAGA	2578
Qy	4019	TTGCTAGACTCCAGTTTGTGTCTTATAGCTCCGGAATAGACTTTTGTGGACGAGTACAC	4078



2579 TTGTTAGACTCCAGTTTGTGTCCTTATAGCCTCCGGAATAGACTTTTGTGACGAGTACAC 2638  
Db  
4079 CAGGCCCAACGAGTAATAGAGAGTACAGCCACCAAGAGTAGTGAATAGACCAATCGGGGCG 4138  
Qy  
2639 CAGGCCCAACGAGTAATAGAGAGTACAGCCACCAAGAGTAGTGAATAGACCAATCGGGGCG 2698  
Db  
4139 GTCAGTAGTCAAGACGCCAACAATAATTCATGACAGGGAACTTTTGTGACATCTTCAGA 4198  
Qy  
2699 GTCAGTAGTCAAGACGCCAACAATAATTCATGACAGGGAACTTTTGTGACATCTTCAGA 2758  
Db  
4199 AAGTTGCTATTACAGTAGTCAATTCGCCGAGCATCAATAATGGGATTATACCAAGAACAC 4258  
Qy  
2759 AAGTTGCTATTACAGTAGTCAATTCGCCGAGCATCAATAATGGGATTATACCAAGAACAC 2818  
Db  
4259 AGTGAAGTCACTCTACCAACTTTGCGGTCTCAGAAAAGCATPAAACAGTTCTACTACC 4318  
Qy  
2819 AGTGAAGTCACTCTACCAACTTTGCGGTCTCAGAAAAGCATPAAACAGTTCTACTACC 2878  
Db  
4319 GCCATTAGTGAACCTTTTCAATCGCCAGTGGAGAGAAAAGGCACAGGACTACTAGC 4378  
Qy  
2879 GCCATTAGTGAACCTTTTCAATCGCCAGTGGAGAGAAAAGGCACAGGACTACTAGC 2938  
Db  
4379 ATTAGCGGGCAAGATGCAACTTTTATCAACAGGGTCTCTATAGATAACCCCTAGCGCTGG 4438  
Qy  
2939 ATTAGCGGGCAAGATGCAACTTTTATCAACAGGGTCTCTATAGATAACCCCTAGCGCTGG 2998  
Db  
4439 GATCATCTTTGGCAACTCTTTCTGCCAATCTAGTCCAAATCACTTATGATACC 4498  
Qy  
2999 GATCATCTTTGGCAACTCTTTCTGCCAATCTAGTCCAAATCACTTATGATACC 3058  
Db  
4499 ATTATTGACAACTTGACAGAGTTGTGATCAGCTCTCTCAAAATGGTCTCTGTAAACGGA 4558  
Qy  
3059 ATTATTGACAACTTGACAGAGTTGTGATCAGCTCTCTCAAAATGGTCTCTGTAAACGGA 3118  
Db  
4559 TGACTCAACTTGCAATPAACTTGAAGCTCAGTCAATGAGTGAACCTGATCAGTTGTG 4618  
Qy  
3119 TGACTCAACTTGCAATPAACTTGAAGCTCAGTCAATGAGTGAACCTGATCAGTTGTG 3178  
Db  
4619 CAGCTGCTCAGCAGATAGGAAACACGGCTTTTCTTACCAAACTCAAGGAATTTATCAA 4678  
Qy  
3179 CAGCTGCTCAGCAGATAGGAAACACGGCTTTTCTTACCAAACTCAAGGAATTTATCAA 3238  
Db  
4679 CTCTGCAACACTTGCGTATGAGGTAGCAAGGAAATGTCTATCTTGAAGTCGACAGTG 4738  
Qy  
3239 CTCTGCAACACTTGCGTATGAGGTAGCAAGGAAATGTCTATCTTGAAGTCGACAGTG 3298  
Db  
4739 AGTGTAGTCTTGAGAAATCTGAGCCGTATTTTATATCAGTGAGTCACTCAGGA 4798  
Qy  
3299 AGTGTAGTCTTGAGAAATCTGAGCCGTATTTTATATCAGTGAGTCACTCAGGA 3358  
Db  
4799 GATCCTCTACGCGGACGATCGTGGCGAATCTGAGGTGCGGATCACCAGCGCCACAGG 4858  
Qy  
3359 GATCCTCTACGCGGACGATCGTGGCGAATCTGAGGTGCGGATCACCAGCGCCACAGG 3406  
Db  
4859 TGCGGTTGCGGCTATATGCGCGACATCAGCGATGGGAAAGATCGGCTCGCCACTT 4918  
Qy  
3407 TGCGGTTGCGGCTATATGCGCGACATCAGCGATGGGAAAGATCGGCTCGCCACTT 3466  
Db  
4919 CGGGCTCATGAGCGCTTGTTCGGCGTGGGTATGGTGGCGAGCCCGCTGGCGGGGACT 4978  
Qy  
3467 CGGGCTCATGAGCGCTTGTTCGGCGTGGGTATGGTGGCGAGCCCGCTGGCGGGGACT 3526  
Db  
4979 GTTGGCGCCATCTCTTTCATGACCAATCTCTTGGCGCGCGGTGTCTCAACGCGCTCAA 5038  
Qy  
3527 GTTGGCGCCATCTCTTTCATGACCAATCTCTTGGCGCGCGGTGTCTCAACGCGCTCAA 3586  
Db  
5039 CTTACTCTGGGCTGCTTCTTAAATGACAGAGTTCGATTAAGGAGAGAGTTCAGTATCTAT 5098  
Qy  
3587 CTTACTCTGGGCTGCTTCTTAAATGACAGAGTTCGATTAAGGAGAGAGTTCAGTATCTAT 3646  
Db  
5099 GATTGGAAGTATGGGAATGGTGAATCCCGCAATCTTCTAGTGTCTTGTAGGTTCTCTATCAG 5158  
Qy  
3647 GATTGGAAGTATGGGAATGGTGAATCCCGCAATCTTCTAGTGTCTTGTAGGTTCTCTATCAG 3706  
Db

5159 ATTATGCCCAACTAAAGCAACCGGAGGAGAGATTTCATGTAAATTTCTCTGACTTTTG 5218  
Qy  
3707 ATTATGCCCAACTAAAGCAACCGGAGGAGAGATTTCATGTAAATTTCTCTGACTTTTG 3766  
Db  
5219 GTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGAGAAATGTCTCTTT 5278  
Qy  
3767 GTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGAGAAATGTCTCTTT 3826  
Db  
5279 GGAGACAGTAATAAGAGTCCCAACAATAAAGAAATCTTGTATTACAGAAACAACTTTCTT 5338  
Qy  
3827 GGAGACAGTAATAAGAGTCCCAACAATAAAGAAATCTTGTATTACAGAAACAACTTTCTT 3886  
Db  
5339 GTTTCGAACCTTTTTCGGTGCCCTTGAACCTATAAATGTAGAGTGTATGTCGGGTAGGAA 5398  
Qy  
3887 GTTTCGAACCTTTTTCGGTGCCCTTGAACCTATAAATGTAGAGTGTATGTCGGGTAGGAA 3946  
Db  
5399 TGGAGCGGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGACT 5458  
Qy  
3947 TGGAGCGGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGACT 4006  
Db  
5459 GATGCCAACTTCAGTGACAAAGTTGCTATTTCGTTCAACCATTCGGAATCCAGAGAAAT 5518  
Qy  
4007 GATGCCAACTTCAGTGACAAAGTTGCTATTTCGTTCAACCATTCGGAATCCAGAGAAAT 4066  
Db  
5519 CAAAGTTGTTGTCTACTATTGATCCAAAGCCAGTGGGTCTTGAACCTGACAAATAGTGTG 5578  
Qy  
4067 CAAAGTTGTTGTCTACTATTGATCCAAAGCCAGTGGGTCTTGAACCTGACAAATAGTGTG 4126  
Db  
5579 CTCGTGTTTGTAGTGCATCTTTGTATGAATAAATCTAGTCTTGTATCTAAATATCTTGA 5638  
Qy  
4127 CTCGTGTTTGTAGTGCATCTTTGTATGAATAAATCTAGTCTTGTATCTAAATATCTTGA 4186  
Db  
5639 CGAGCAAGCGGATAAATACCAAACTTAAACCTCTTTTAAACCTTAAAGGCAAGTA 5698  
Qy  
4187 CGAGCAAGCGGATAAATACCAAACTTAAACCTCTTTTAAACCTTAAAGGCAAGTA 4246  
Db  
5699 TCTCTGCTGTATTAAACCCCAATCAGTCTGATCTGATCCTCATCACTTGAAGGGG 5758  
Qy  
4247 TCTCTGCTGTATTAAACCCCAATCAGTCTGATCCTCATCACTTGAAGGGG 4306  
Db  
5759 ACTATCTGTTTGTAGAGAAATTTGCGAGATGCGATATCGAGAAAGAGTACGCTGATT 5818  
Qy  
4307 ACTATCTGTTTGTAGAGAAATTTGCGAGATGCGATATCGAGAAAGAGTACGCTGATT 4366  
Db  
5819 TAAACGTGAAATTTTATCTCAAGATCTCTGCTCGCGGTTTTCGGGTGATGACGGTGAAC 5878  
Qy  
4367 TAAACGTGAAATTTTATCTCAAGATCTCTGCTCGCGGTTTTCGGGTGATGACGGTGAAC 4426  
Db  
5879 CTCTGACATCAGCTCCGAGAGCGTCAAGTCTCTGTAAGCGGATGCGGGAGC 5938  
Qy  
4427 CTCTGACATCAGCTCCGAGAGCGTCAAGTCTCTGTAAGCGGATGCGGGAGC 4486  
Db  
5939 AGACAAGCCCGTCAGGCGCGTCAAGGCTGTGCGGGTGTTCGGGCGCAGCCATGACC 5998  
Qy  
4487 AGACAAGCCCGTCAGGCGCGTCAAGGCTGTGCGGGTGTTCGGGCGCAGCCATGACC 4546  
Db  
5999 CAGTCACTGAGGATGAGGAGTGTATCTGCTTAACTATGCGGCTCAGAGCAGATTG 6058  
Qy  
4547 CAGTCACTGAGGATGAGGAGTGTATCTGCTTAACTATGCGGCTCAGAGCAGATTG 4606  
Db  
6059 TACTGAGAGTGCACCATATGCGGTGTAATAACCCACAGATGCGTAAGGAGAAATACC 6118  
Qy  
4607 TACTGAGAGTGCACCATATGCGGTGTAATAACCCACAGATGCGTAAGGAGAAATACC 4666  
Db  
6119 GCATCAGGCGCTCTTTCGGCTTCCCTCGCTCACTGATCTGCTGCGCTCGGTTGGCTGC 6178  
Qy  
4667 GCATCAGGCGCTCTTTCGGCTTCCCTCGCTCACTGATCTGCTGCGCTCGGTTGGCTGC 4726  
Db  
6179 GCGAGCGGTATCAGTCACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATA 6238  
Qy  
4727 GCGAGCGGTATCAGTCACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATA 4786  
Db

QY 6239 ACGCAGGAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGAGCCG 6298  
Db 4787 ACGCAGGAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGAGCCG 4846  
QY 6299 CGTTGTCGCGGTTTTTCCATAGGCTCCGCCGCCCTCGAGAGCATCACAAAATCGAGCCT 6358  
Db 4847 CGTTGTCGCGGTTTTTCCATAGGCTCCGCCGCCCTCGAGAGCATCACAAAATCGAGCCT 4906  
QY 6359 CAAATCAGAGTGGGGAACCCGACAGAGCATATAAGATACAGAGGTTTCCCGCTGGAA 6418  
Db 4907 CAAATCAGAGTGGGGAACCCGACAGAGCATATAAGATACAGAGGTTTCCCGCTGGAA 4966  
QY 6419 GCTCCCTCGTGGGCTCTCTGTTTCGAGCCTCGCCTTACCGGATACCTGTCCGCGCTTC 6478  
Db 4967 GCTCCCTCGTGGGCTCTCTGTTTCGAGCCTCGCCTTACCGGATACCTGTCCGCGCTTC 5026  
QY 6479 TCCCTTCGGAAGAGTGGGCTCTCTGTTTCGAGCCTCGCCTTACCGGATACCTGTCCGCGCTTC 6538  
Db 5027 TCCCTTCGGAAGAGTGGGCTCTCTGTTTCGAGCCTCGCCTTACCGGATACCTGTCCGCGCTTC 5086  
QY 6539 AGTCTGTTCCGCTCCAGCTGGGCTGTGACGACCCCGGTTACCGCCGACCGCTGCG 6598  
Db 5087 AGTCTGTTCCGCTCCAGCTGGGCTGTGACGACCCCGGTTACCGCCGACCGCTGCG 5146  
QY 6599 CTTTATCCGGTAACCTATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGG 6658  
Db 5147 CTTTATCCGGTAACCTATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGG 5206  
QY 6659 CAGCAGCCACTGGTAACAGGATACAGAGCAGGATATGAGCGGCTGCTACAGAGTTCT 6718  
Db 5207 CAGCAGCCACTGGTAACAGGATACAGAGCAGGATATGAGCGGCTGCTACAGAGTTCT 5266  
QY 6719 TGAAGTGGTGGCTAACTACGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGC 6778  
Db 5267 TGAAGTGGTGGCTAACTACGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGC 5326  
QY 6779 TGAAGCAGTACCTTCGGAAGAGTGGTATGCTTTGATTCGGGCAAAACAAACCCAGC 6838  
Db 5327 TGAAGCAGTACCTTCGGAAGAGTGGTATGCTTTGATTCGGGCAAAACAAACCCAGC 5386  
QY 6839 CTGCTAGCGGTGGTTTTTTGTTTCAAGCAGCAGATACCGCAGAGAAAGAGATCTC 6898  
Db 5387 CTGCTAGCGGTGGTTTTTTGTTTCAAGCAGCAGATACCGCAGAGAAAGAGATCTC 5446  
QY 6899 AAGAAGATCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAAAGAAACCTCAGCTT 6958  
Db 5447 AAGAAGATCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAAAGAAACCTCAGCTT 5506  
QY 6959 AAGGATTTTGTCTAGATATCAAAAGGATCTTCACTAGATCTTTTAAATTTAA 7018  
Db 5507 AAGGATTTTGTCTAGATATCAAAAGGATCTTCACTAGATCTTTTAAATTTAA 5566  
QY 7019 AATGAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGGTCTGACAGTTACCAAT 7078  
Db 5567 AATGAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGGTCTGACAGTTACCAAT 5626  
QY 7079 GCTTAATCAGTGAGGACCTATCTCAGGATCTGCTATTTGTTTCAATCCATAGTTCCT 7138  
Db 5627 GCTTAATCAGTGAGGACCTATCTCAGGATCTGCTATTTGTTTCAATCCATAGTTCCT 5686  
QY 7139 GACTCCCGTCTGTAGATAACTACGATACGGAGGGGTTACCACTCGCCCGCAGTGTG 7198  
Db 5687 GACTCCCGTCTGTAGATAACTACGATACGGAGGGGTTACCACTCGCCCGCAGTGTG 5746  
QY 7199 CAATGATACCGGACCCACGCTCACCGGCTCCAGATTTATCAAGCAATAAACAGCCAG 7258  
Db 5747 CAATGATACCGGACCCACGCTCACCGGCTCCAGATTTATCAAGCAATAAACAGCCAG 5806  
QY 7259 CCGGAGGGCCGAGCGAGTGGTCTGCACTTTATCCGCTCCCTCCAGTCTATTA 7318  
Db 5807 CCGGAGGGCCGAGCGAGTGGTCTGCACTTTATCCGCTCCCTCCAGTCTATTA 5866  
QY 7319 ATTGTTGCCGGAAGTAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTTG 7378

Db 5867 ATTGTTGCCGGAAGCTAGAGTAAGTAGTTCCGAGTTAATAGTTTGCACAAAGTTGTTG 5926  
QY 7379 CCATTGCTGAGGACATCGTGTCTCAGCTCGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 7438  
Db 5927 CCATTGCTGAGGACATCGTGTCTCAGCTCGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5986  
QY 7439 GTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCT 7498  
Db 5987 GTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCT 6046  
QY 7499 CTTTCGCTCTCCGATCGTTTGTGAGAGTAAGTTGCGCGCAGTGTATCACTCATGTTA 7558  
Db 6047 CTTTCGCTCTCCGATCGTTTGTGAGAGTAAGTTGCGCGCAGTGTATCACTCATGTTA 6106  
QY 7559 TGGCAGCAGTGCATTAATCTCTTACTGTCTATGTCATCCGTAAGATGCTTTTCTGTGACTG 7618  
Db 6107 TGGCAGCAGTGCATTAATCTCTTACTGTCTATGTCATCCGTAAGATGCTTTTCTGTGACTG 6166  
QY 7619 GTGAGTACTCAACCAAGTCAATCTGAGATAGTGTATGCGGACCGAGTGTCTTTGCC 7678  
Db 6167 GTGAGTACTCAACCAAGTCAATCTGAGATAGTGTATGCGGACCGAGTGTCTTTGCC 6226  
QY 7679 CGCGCTCAACACGGAATAATACCGCGCCACATAGCAAACTTTTAAAGTGTCTCATCTG 7738  
Db 6227 CGCGCTCAACACGGAATAATACCGCGCCACATAGCAAACTTTTAAAGTGTCTCATCTG 6286  
QY 7739 GAAACGTTCTTCGCGCGGAAACTCTCAGGATCTTACCGTGTGAGATCCAGTTCGA 7798  
Db 6287 GAAACGTTCTTCGCGCGGAAACTCTCAGGATCTTACCGTGTGAGATCCAGTTCGA 6346  
QY 7799 TGTAAACCCACTCGTGCACCCCACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTG 7858  
Db 6347 TGTAAACCCACTCGTGCACCCCACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTG 6406  
QY 7859 GGTGAGCAAAAACAGGAAGCAAAATGCGCAAAAAGGGAATAAGGGGAGACAGGAAT 7918  
Db 6407 GGTGAGCAAAAACAGGAAGCAAAATGCGCAAAAAGGGAATAAGGGGAGACAGGAAT 6466  
QY 7919 GTTGAATACCTACTACTCTTCTCTTTTCAATATTATTGAAAGCTTTATCAGGGTTATTGTC 7978  
Db 6467 GTTGAATACCTACTACTCTTCTCTTTTCAATATTATTGAAAGCTTTATCAGGGTTATTGTC 6526  
QY 7979 TCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAACAAATAGGGGTTCCGCGCA 8038  
Db 6527 TCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAACAAATAGGGGTTCCGCGCA 6586  
QY 8039 CATTTCCCGAAAGTGCCACCTGAGCTCTAGAAACCAATTTATCATGACATTAACCT 8098  
Db 6587 CATTTCCCGAAAGTGCCACCTGAGCTCTAGAAACCAATTTATCATGACATTAACCT 6646  
QY 8099 ATAAAAATAGCGGTATCAGAGGCCCTTTTCGCTCTTCAAGAAATTAATCTCATGTTTGACA 8158  
Db 6647 ATAAAAATAGCGGTATCAGAGGCCCTTTTCGCTCTTCAAGAAATTAATCTCATGTTTGACA 6706  
QY 8159 GCTTATCATGATAAGCTGACTCATGTTGGTATTGTGAATAGCAGATCCGGGAACAC 8218  
Db 6707 GCTTATCATGATAAGCTGACTCATGTTGGTATTGTGAATAGCAGATCCGGGAACAC 6766  
QY 8219 TGAATAATAACAGTTATTATTCG 8241  
Db 6767 TGAATAATAACAGTTATTATTCG 6789

RESULT 5  
PPP2IC3  
LOCUS  
DEFINITION P.pastoris DNA for pPIC3 expression vector.  
ACCESSION Z46231  
VERSION Z46231.1 GI:559509  
KEYWORDS beta-lactamase.  
SOURCE Pichia pastoris  
ORGANISM Pichia pastoris

7757 bp DNA linear PLN 22-OCT-1994

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; Pichia.  
 1 (bases 1 to 7757)  
 Scorer, C.A., Clare, J.J., McCombie, W.R., Romanos, M.A. and  
 Sreekrishna, K.  
 Rapid selection using G418 of high copy number transformants of  
 Pichia pastoris for high-level foreign gene expression  
 Biotechnology (N.Y.) 12 (2), 181-184 (1994)  
 94153511  
 7764433  
 2 (bases 1 to 7757)  
 Romanos, M.A.  
 Direct Submission  
 Submitted (18-OCT-1994) Romanos M. A., Wellcome Research  
 Laboratories, Biology Division, Langley Court, Beckenham, Kent, UK,  
 BR3 3BS

FEATURES  
 source  
 1. .7757  
 /organism="Pichia pastoris"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4922"  
 /clone="pPIC3"  
 1. .7757  
 /standard\_name="pPIC3"  
 /function="Pichia intracellular expression vector"  
 1. .354  
 /standard\_name="AOX1"  
 complement(1237..3267)  
 /standard\_name="HIS4"  
 3589..4036  
 /standard\_name="3' AOX1"  
 complement(5606..6466)  
 /product="beta-lactamase"  
 /standard\_name="AmpR"  
 6766..7703  
 /standard\_name="AOX1"

misc\_feature  
 1. .7757  
 /function="Pichia intracellular expression vector"

terminator  
 1. .354  
 /standard\_name="AOX1"  
 complement(1237..3267)  
 /standard\_name="HIS4"  
 3589..4036  
 /standard\_name="3' AOX1"  
 complement(5606..6466)  
 /product="beta-lactamase"  
 /standard\_name="AmpR"  
 6766..7703  
 /standard\_name="AOX1"

mRNA  
 1. .7757  
 /standard\_name="AOX1"

misc\_feature  
 1. .7757  
 /function="Pichia intracellular expression vector"

mRNA  
 1. .7757  
 /standard\_name="AOX1"

promoter  
 1. .7757  
 /standard\_name="AOX1"

ORIGIN  
 Query Match 81.5%; Score 6713.6; DB 8; Length 7757;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 6771; Conservative 0; Mismatches 4; Indels 15; Gaps 4;  
 1453 AATT CGCCTTAGACATGACTTTCTCAGTTCAGAGTGGGCACCTACGAGAACCGGTC 1512  
 Db 1 AATT CGCCTTAGACATGACTTTCTCAGTTCAGAGTGGGCACCTACGAGAACCGGTC 60  
 1513 TTGCTAGATTCCTATCAGAGGATGTCAGATGCCATTTGCTCAGAGATGCAGGCTTCA 1572  
 QY 61 TTGCTAGATTCCTATCAGAGGATGTCAGATGCCATTTGCTCAGAGATGCAGGCTTCA 120  
 Db 173 TTTTGTGACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTGTTT 1632  
 QY 121 TTTTGTGACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTGTTT 180  
 Db 1633 CTTCTCGTACAGAGTTGCTCTGATCAGCCTATCTCCGAGCTGATGAATCTTGTGCTGA 1692  
 QY 181 CTTCTCGTACAGAGTTGCTCTGATCAGCCTATCTCCGAGCTGATGAATCTTGTGCTGA 240  
 Db 1693 GGGGTTTCGGAAATCAATTCAGATTGATGTTTTTTCTTGGTATTTCCACCTCTCTTCAG 1752  
 QY 241 GGGGTTTCGGAAATCAATTCAGATTGATGTTTTTTCTTGGTATTTCCACCTCTCTTCAG 300  
 Db 1753 AGTACAGAAGATTAGTGAGAAGTTCTGTTTGTGCAAGCTTATCGATAGCTTTAATGCGG 1812  
 QY 301 AGTACAGAAGATTAGTGAGAAGTTCTGTTTGTGCAAGCTTATCGATAGCTTTAATGCGG 360  
 Db 1813 TAGTTTATCAGAGTTAAATTCGTAAACGAGTCAGGCACCGTGTATGAATCTTAAATGCG 1872  
 QY 361 TAGTTTATCAGAGTTAAATTCGTAAACGAGTCAGGCACCGTGTATGAATCTTAAATGCG 420  
 Db 1873 GCTCATCGTCATCTCGGACCGTCAACCGCTGATGCTGTAGGCATAGCGTTGGTTATGCC 1932  
 QY 421 GCTCATCGTCATCTCGGACCGTCAACCGCTGATGCTGTAGGCATAGCGTTGGTTATGCC 480  
 Db



Db 3707 AAAGCAACCGGAGGAGGATTTTCATGGTAAATTTCTCTGACTTTTGGTCATCAGTAGAC 3766  
 QY 5232 TCAGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCTCTTTGGAGACAGTAAAT 5291  
 Db 3767 TCAGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCTCTTTGGAGACAGTAAAT 3826  
 QY 5292 GAAGTCCCAACCAATAAGAAATCTCTGTTATCAGGAACAAATCTCTCTTTTCGAACTTTT 5351  
 Db 3827 GAAGTCCCAACCAATAAGAAATCTCTGTTATCAGGAACAAATCTCTCTTTTCGAACTTTT 3886  
 QY 5352 TCGGTCCCTTGAACTATAAATAATGTAGAGTGNATGTCTGGGTAGAAATGAGACGGGCAAA 5411  
 Db 3887 TCGGTCCCTTGAACTATAAATAATGTAGAGTGNATGTCTGGGTAGAAATGAGACGGGCAAA 3946  
 QY 5412 TGCTTACCTTCGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGATGCCAACTTCA 5471  
 Db 3947 TGCTTACCTTCGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGATGCCAACTTCA 4006  
 QY 5472 GTGACAAAGTGTCTATTTTCGTTCAACCAATTCGGAATCCAGAGAAATCAAAAGTTGTTGT 5531  
 Db 4007 GTGACAAAGTGTCTATTTTCGTTCAACCAATTCGGAATCCAGAGAAATCAAAAGTTGTTGT 4066  
 QY 5532 CTACTATTGATCCAAAGCAGTGGGTCTTGAACCTGCAATAGTGTCTCGTGTGTTGAG 5591  
 Db 4067 CTACTATTGATCCAAAGCAGTGGGTCTTGAACCTGCAATAGTGTCTCGTGTGTTGAG 4126  
 QY 5592 GTCATCTTTGTATGAATAAATCTAGTCTTGAATCTTGAATCTTGAAGTGTCTGCTGTTGAT 5651  
 Db 4127 GTCATCTTTGTATGAATAAATCTAGTCTTGAATCTTGAATCTTGAAGTGTCTGCTGTTGAT 4186  
 QY 5652 TAAATACCAAAATCTTAAATCTTTTAAACCTTAAAGGACAAAGTATGCTGCTGCTGAT 5711  
 Db 4187 TAAATACCAAAATCTTAAATCTTTTAAACCTTAAAGGACAAAGTATGCTGCTGCTGAT 4246  
 QY 5712 TAAACCCCAAAATCAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 5771  
 Db 4247 TAAACCCCAAAATCAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4306  
 QY 5772 AGAGAAATTTGGGAGATCGGATATCGAGAAAGAGTACGCTGATTTTAAACGTTGAAT 5831  
 Db 4307 AGAGAAATTTGGGAGATCGGATATCGAGAAAGAGTACGCTGATTTTAAACGTTGAAT 4366  
 QY 5832 TATCTCAAGATCTCTGCTCGGCTGTTGCGGTATGACGGTGAAGAACTCTGACACATGC 5891  
 Db 4367 TATCTCAAGATCTCTGCTCGGCTGTTGCGGTATGACGGTGAAGAACTCTGACACATGC 4426  
 QY 5892 AGCTCCGAGACGGTACAGCTTCTGTTAAGCGGATGCGGAGCAGACAAAGCCGCTC 5951  
 Db 4427 AGCTCCGAGACGGTACAGCTTCTGTTAAGCGGATGCGGAGCAGACAAAGCCGCTC 4486  
 QY 5952 AGGCGCGTACGCGGTGTTGCGGCTGTTGCGGCTGTTGCGGCTGTTGCGGCTGTTGCGGCTGTT 6011  
 Db 4487 AGGCGCGTACGCGGTGTTGCGGCTGTTGCGGCTGTTGCGGCTGTTGCGGCTGTTGCGGCTGTT 4546  
 QY 6012 ATAGCGGAGTGTATCTGCTTAACTATGCGGATCAGAGCAGATTTGACTGAGAGTGCA 6071  
 Db 4547 ATAGCGGAGTGTATCTGCTTAACTATGCGGATCAGAGCAGATTTGACTGAGAGTGCA 4606  
 QY 6072 CCATATGCGGTGTGAATAACGCAAGATGCGTGAAGGAGAAATACCGCATCAGGCGCTC 6131  
 Db 4607 CCATATGCGGTGTGAATAACGCAAGATGCGTGAAGGAGAAATACCGCATCAGGCGCTC 4666  
 QY 6132 TTCGCTTCTCGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6191  
 Db 4667 TTCGCTTCTCGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4726  
 QY 6192 AGTCACTCAAAAGGCGGTATACGTTATCAAGATCAGGGGATTAACGAGGAAGAA 6251  
 Db 4727 AGTCACTCAAAAGGCGGTATACGTTATCAAGATCAGGGGATTAACGAGGAAGAA 4786  
 QY 6252 CATGTAGCAAAAGGCGGTATACGTTATCAAGATCAGGGGATTAACGAGGAAGAA 6311  
 Db 4787 CATGTAGCAAAAGGCGGTATACGTTATCAAGATCAGGGGATTAACGAGGAAGAA 4846

QY 6312 TTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTG 6371  
 Db 4847 TTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTG 4906  
 QY 6372 GCGAAACCCGACAGGACTATAAAGATACCAAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG 6431  
 Db 4907 GCGAAACCCGACAGGACTATAAAGATACCAAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG 4966  
 QY 6432 CTCCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAG 6491  
 Db 4967 CTCCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAG 5026  
 QY 6492 CGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTTAGGTCCGTTCCGCTC 6551  
 Db 5027 CGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTTAGGTCCGTTCCGCTC 5086  
 QY 6552 CAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGTTAA 6611  
 Db 5087 CAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGTTAA 5146  
 QY 6612 CTATCGTCTTGAGTCCAAACCCGTTAAGACACGACTTATCGCCACTGCGACGACCACTGG 6671  
 Db 5147 CTATCGTCTTGAGTCCAAACCCGTTAAGACACGACTTATCGCCACTGCGACGACCACTGG 5206  
 QY 6672 TAAAGGATTAAGACAGGAGGTATGTAGGGGTGCTACAGAGTTCTTGAAGTGGTGGC 6731  
 Db 5207 TAAAGGATTAAGACAGGAGGTATGTAGGGGTGCTACAGAGTTCTTGAAGTGGTGGC 5266  
 QY 6732 TAACTACGCTTACACTAGAGGACAGTATTTGGTATCTGCGCTCTCTGTAAGCAGTTAC 6791  
 Db 5267 TAACTACGCTTACACTAGAGGACAGTATTTGGTATCTGCGCTCTCTGTAAGCAGTTAC 5326  
 QY 6792 CTTCCGAAAAAGAGTGTGTTAGTCTTGTATCCGGAACAAACACCGCTGTTAGCGGTGG 6851  
 Db 5327 CTTCCGAAAAAGAGTGTGTTAGTCTTGTATCCGGAACAAACACCGCTGTTAGCGGTGG 5386  
 QY 6852 TTTTCTTTGTTTGAAGCAGAGATTTAGCGCAGAAAAAAGGATCTCAAGAAGATCCTTT 6911  
 Db 5387 TTTTCTTTGTTTGAAGCAGAGATTTAGCGCAGAAAAAAGGATCTCAAGAAGATCCTTT 5446  
 QY 6912 GATCTTTTCTACGGGTCTGACGCTCAGTGAACAGAAAACTCAGTTAAGGATTTTGT 6971  
 Db 5447 GATCTTTTCTACGGGTCTGACGCTCAGTGAACAGAAAACTCAGTTAAGGATTTTGT 5506  
 QY 6972 CATGAGATTATCAAAAAGGATCTTACCTAGATCTCTTTAAATTAATAAATGAAGTTTAA 7031  
 Db 5507 CATGAGATTATCAAAAAGGATCTTACCTAGATCTCTTTAAATTAATAAATGAAGTTTAA 5566  
 QY 7032 ATCAATCTAAAGTATATAGGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTGA 7091  
 Db 5567 ATCAATCTAAAGTATATAGGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTGA 5626  
 QY 7092 GGCACCTATCTCAGCGATCTGTTATTTTCTTATTCATCATAGTTGCTGACTCCCGCTCGT 7151  
 Db 5627 GGCACCTATCTCAGCGATCTGTTATTTCTTATTCATCATAGTTGCTGACTCCCGCTCGT 5686  
 QY 7152 GTAGATACTACGATACGGGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGG 7211  
 Db 5687 GTAGATACTACGATACGGGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGG 5746  
 QY 7212 AGACCCAGCTTACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCGGAGGCGCA 7271  
 Db 5747 AGACCCAGCTTACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCGGAGGCGCA 5806  
 QY 7272 GCGCAGAAAGTGTCTCGAACTTTATCCGCTCCATCCAGTCTTATTAATGTTGCGGGA 7331  
 Db 5807 GCGCAGAAAGTGTCTCGAACTTTATCCGCTCCATCCAGTCTTATTAATGTTGCGGGA 5866  
 QY 7332 AGCTAGAGTAAAGTGTGCGCAGTTAATAGTTTGGCAGAACGTTGTTGCCATGCTGTCAG 7391  
 Db 5867 AGCTAGAGTAAAGTGTGCGCAGTTAATAGTTTGGCAGAACGTTGTTGCCATGCTGTCAG 5926









2927 TAAGGTAACCATCTTCTTAGTGGAGATGCAAAAACATTTCTTGCACACGCAACTT 2986  
3061 TGGCAGGAACACCCAGCATCAGGAAGTGGGAAGCAGAAATTCGGTTCCACCGAATAT 3120  
2987 TGGCAGGAACACCCAGCATCAGGAAGTGGGAAGCAGAAATTCGGTTCCACCGAATAT 3046  
3121 AGAGCCAACTTCTCAATAGGTTCTCAAAACGAGAGCAGACTACACCGAGGCAAGTCT 3180  
3047 AGAGCCAACTTCTCAATAGGTTCTCAAAACGAGAGCAGACTACACCGAGGCAAGTCT 3106  
3181 CAATCTGCAACGCTCTCCGTTAGTTGAGCTTCAATGGAATTTCCCTGACGTTATCTATAGAGA 3240  
3107 CAATCTGCAACGCTCTCCGTTAGTTGAGCTTCAATGGAATTTCCCTGACGTTATCTATAGAGA 3166  
3241 GATCAATGGCTCTCTTAACGTTATCTGCAATTCGATAAGTTCTCTGGGAAGGAGCTT 3300  
3167 GATCAATGGCTCTCTTAACGTTATCTGCAATTCGATAAGTTCTCTGGGAAGGAGCTT 3226  
3301 CTAACACAGGTTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTT 3360  
3227 CTAACACAGGTTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTT 3286  
3361 TGTCAACATTTGAGCAACATGTGCACAATGGTTTGAATTAATCCAAATCTGTTCCG 3420  
3287 TGTCAACATTTGAGCAACATGTGCACAATGGTTTGAATTAATCCAAATCTGTTCCG 3346  
3421 TTTCTGATAGGAGCAGAGGAGCATCTCAATTTCTGTGAGAGGCTTGAAGACGT 3480  
3347 TTTCTGATAGGAGCAGAGGAGCATCTCAATTTCTGTGAGAGGCTTGAAGACGT 3406  
3481 CAATTTTGCACAATTAATAGCACTTCAGAAAGGACTCTTTTAGTTTGAATTTCTTCT 3540  
3407 CAATTTTGCACAATTAATAGCACTTCAGAAAGGACTCTTTTAGTTTGAATTTCTTCT 3466  
3541 TAGTTGTTCTTCTGTTGATCTGCTGTTGGATCTCTCTTCTCTAGTACCTTTAGGG 3600  
3467 TAGTTGTTCTTCTGTTGATCTGCTGTTGGATCTCTCTTCTCTAGTACCTTTAGGG 3526  
3601 ACTTCATATCCAGGTTCTCTCCACTCGTCCAACTGTCACACGCTTGGCACAATCTAA 3660  
3527 ACTTCATATCCAGGTTCTCTCCACTCGTCCAACTGTCACACGCTTGGCACAATCTAA 3586  
3661 CTAATGCAAAATAAATAAGTACAGACATTCACAGCTATATCTTCTTGAATTTAGCTT 3720  
3587 CTAATGCAAAATAAATAAGTACAGACATTCACAGCTATATCTTCTTGAATTTAGCTT 3646  
3721 CTGCAAGTTTCATCAGCTTCTCTCCCTAAATTTTAGCGTTCAACAAACTTCGTCTCAATA 3780  
3647 CTGCAAGTTTCATCAGCTTCTCTCCCTAAATTTTAGCGTTCAACAAACTTCGTCTCAATA 3706  
3781 ACCGTTTGGTATAAGAACCTTCTGGAGCATTCGTCTTACGATCCCAAGGTTGCTTCCA 3840  
3707 ACCGTTTGGTATAAGAACCTTCTGGAGCATTCGTCTTACGATCCCAAGGTTGCTTCCA 3766  
3841 TGGCTCTAGACCTTTGATTTGGCCAAACAGAGAGTGCGTTCCCAAGTGACAGAAACCAA 3900  
3767 TGGCTCTAGACCTTTGATTTGGCCAAACAGAGAGTGCGTTCCCAAGTGACAGAAACCAA 3826  
3901 CACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3960  
3827 CACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3886  
3961 GCAACTTTTGTGTTGCTCCGATGTAGCAGCTTTATACCAAAACCGTGACAGAGATT 4020  
3887 GCAACTTTTGTGTTGCTCCAGATGTAGCAGCTTTATACCAAAACCGTGACAGAGATT 3946  
4021 GGTAGACTCCAGTTTGTGTTCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACCA 4080  
3947 GGTAGACTCCAGTTTGTGTTCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACCA 4006  
4081 GGCCCAACGAGTAAATAGAGAGTACAGCCACCAAAAGTGTGAATAGACCATCGGGCGGT 4140

4007 GGCCCAACGAGTAATTTAGAAAGTCAAGCCACCAAAAGTGTGAATAGACCATCGGGCGGT 4066  
4141 CAGTAGTCAAAACAGCCCAACAAAATTTCACTCACAAGGAACTTTTGTGACATCTTTCAGAAA 4200  
4067 CAGTAGTCAAAACAGCCCAACAAAATTTCACTCACAAGGAACTTTTGTGACATCTTTCAGAAA 4126  
4201 GTTCTGTTTTCAGTAGTCAATTCGAGCATCAATTAATGGGGATTATACAGAAAGCAACAG 4260  
4127 GTTCTGTTTTCAGTAGTCAATTCGAGCATCAATTAATGGGGATTATACAGAAAGCAACAG 4186  
4261 TGGAGTCACTACTACCAACTTTGGGCTCTCAGAAAAAGCATAAAACAGTTCTACTACCGC 4320  
4187 TGGAGTCACTACTACCAACTTTGGGCTCTCAGAAAAAGCATAAAACAGTTCTACTACCGC 4246  
4321 CATTAGTGAACCTTTTCAAAATCCGAGTGGAGAGAAAAAGGCAAGGCAATCTAGCAT 4380  
4247 CATTAGTGAACCTTTTCAAAATCCGAGTGGAGAGAAAAAGGCAAGGCAATCTAGCAT 4306  
4381 TAGCGGCAAGGATGCAACTTTTATCAACAGAGTCTCTATAGATAACCTTAGCGCTTGGGA 4440  
4307 TAGCGGCAAGGATGCAACTTTTATCAACAGAGTCTCTATAGATAACCTTAGCGCTTGGGA 4366  
4441 TCATCTTTTGGACAACTCTTTCTGCCAAATCTAGGTTCCAAATCACTTCATTTGATACCAT 4500  
4367 TCATCTTTTGGACAACTCTTTCTGCCAAATCTAGGTTCCAAATCACTTCATTTGATACCAT 4426  
4501 TATTGTACAACTTGAGCAAGTTGTGATCAGCTCTCAAAATTTGGTCTCTGTAAACGGATG 4560  
4427 TATTGTACAACTTGAGCAAGTTGTGATCAGCTCTCTCAAAATTTGGTCTCTGTAAACGGATG 4486  
4561 ACTCAACTTGCACATTAACCTTTGAAGCTCAGTTCGATGAGTGAACCTTGAATTCAGTTGTGCA 4620  
4487 ACTCAACTTGCACATTAACCTTTGAAGCTCAGTTCGATGAGTGAACCTTGAATTCAGTTGTGCA 4546  
4621 GCTGGTCAAGCATAGGGAACACACGGCTTTTCTCAACCAACTCAAGGAATTTCAAACT 4680  
4547 GCTGGTCAAGCATAGGGAACACACGGCTTTTCTCAACCAACTCAAGGAATTTCAAACT 4606  
4681 CTGCAACACTTGGTATGAGTAGCAAGGAAATGTCTACTTTGAAGTTCGAGCAGTGAG 4740  
4607 CTGCAACACTTGGTATGAGTAGCAAGGAAATGTCTACTTTGAAGTTCGAGCAGTGAG 4666  
4741 TGAGTCTTTCAGAAATTTCTGAAGCCGTATTTTATATCAGTGAAGTCAATCAGGATG 4800  
4667 TGAGTCTTTCAGAAATTTCTGAAGCCGTATTTTATATCAGTGAAGTCAATCAGGATG 4726  
4801 TCCTCTAGCCGCAAGCATCTGTCGCGA ----- 4828  
4727 TCCTCTAGCCGCAAGCATCTGTCGCGA ----- 4786  
4829 ----- 4828  
4787 CTGCTGAAGAAGTGTGTGCTGACTCATACCAGGCTGAATCGCCCATCATCCAGCCAGA 4846  
4829 ----- 4828  
4847 AAGTGAAGGAGCCAGGTTGATGAGAGCTTTGTTGTAGTGGACAGTTGGTGATTTTGA 4906  
4829 ----- 4828  
4907 ACTTTTGTGTTGCCACGGAACGGTCTGCGTTGTGCGGAAGATGCGTGATCTGATCTTCA 4966  
4829 ----- 4828  
4967 ACTCAGCAAAAGTTTCGATTTTATCAACAAAGCCCGTCCCGTCAAGTCAAGCGTAATGCT 5026  
4829 ----- 4828  
5027 CTGCGAGTGTTCACCAACCAATTAACCAATTCGATTAGAAAAAATCATCTGAGCATCAAAATG 5086  
4829 ----- 4828  
5087 AAACGTCAATTTTATTCATATCAGGATTAATCAATACCATATTTTGTAAAAAGCCGTTTCTG 5146

QY 4829 ----- 4828  
Db 5147 TAATGAAGAGAGAAACTCACCGAGGCACTTCATAGGATGCAAGATCCTGGTATCGGTC 5206  
QY 4829 ----- 4828  
Db 5207 TCGGATCCGACTCGTCAACATCAATACAACTATTAAATTCCTCCCTCGTCAAAATAAG 5266  
QY 4829 ----- 4828  
Db 5267 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGCAAAAGCTT 5326  
QY 4829 ----- 4828  
Db 5327 ATGCATTTCTTTCCAGACTTGTTCACACAGGCCAGCCATTACGTCGTGTCATCAAAATCACT 5386  
QY 4829 ----- 4828  
Db 5387 CGCATCAACCAACCGTTATTTCATTTCGTGATTGCGCTGACGAGACGAATACGCGATC 5446  
QY 4829 ----- 4828  
Db 5447 GCTGTTAAAGGACAAATTACAAACAGGAATCGAATGCAACCGCGCGAGGAACACTGCCAG 5506  
QY 4829 ----- 4828  
Db 5507 CGCATCAACAAATTTTCACTCGAATCAGGATATTCTTCTAATACCTGGAAATGCTGTTTT 5566  
QY 4829 ----- 4828  
Db 5567 CCGGGGATCGCAGTGTGAGTAACCATGTCATCATCAGGAGTACGGATATAATGCTTGAT 5626  
QY 4829 ----- 4828  
Db 5627 GGTGGAGAGGCAATAAATTCGTGAGCCAGTTTGTAGTCTGACCATCTCATCTGTGAAATC 5686  
QY 4829 ----- 4828  
Db 5687 ATTGCAACGCTACTCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGCTTCCCAT 5746  
QY 4829 ----- 4828  
Db 5747 CAATCGATAGATGTGCACTGATGCGCGACATTATCGCGACCCATTATATACCCATA 5806  
QY 4829 ----- 4828  
Db 5807 TAAATCAGCATCCATGTTGGAAATTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5866  
QY 4829 ----- 4828  
Db 5867 ATGGCTCATAACACCCCTTGATTACTGTTTATGTAAGCAGACAGTTTATTGTTTATGA 5926  
QY 4829 ----- 4828  
Db 5927 TGATATATTTTATCTGTGCAATGTAAATCATAGAGATTTTGAGACACAACGTTGGCTTTC 5986  
QY 4829 ----- CTTGCAAGTTCGCATCACCGCGCCACAGGTGCGGTGCTGCGCGCTATAT 4879  
Db 5987 CCCCCCCCCCTGCAAGTTCGCATCACCGCGCCACAGGTGCGGTGCTGCGCGCTATAT 6046  
QY 4880 CGCGACATCACCGATGGGAGATCGGCTCGCCACTTCGGGCTCATGAGCGCTGTTT 4939  
Db 6047 CGCGACATCACCGATGGGAGATCGGCTCGCCACTTCGGGCTCATGAGCGCTGTTT 6106  
QY 4940 CGGCTGCGGTATGTGGCAGCCCGCTGGCGCGGAGCTGTTGGCGCCCATCTCTTTCGA 4999  
Db 6107 CGGCTGCGGTATGTGGCAGCCCGCTGGCGCGGAGCTGTTGGCGCCCATCTCTTTCGA 6166  
QY 5000 TGCACCATTCCTTGGCGCGGCGGTGCTCAACGGCTCAACCTACTACTGGCTGCTTCCT 5059  
Db 6167 TGCACCATTCCTTGGCGCGGCGGTGCTCAACGGCTCAACCTACTACTGGCTGCTTCCT 6226

QY 5060 AATGCAGAGTCCGATTAAGGAGAGCGTGCAGTATCTATGATTGGAAGTATGGGAATGTT 5119  
Db 6227 AATGCAGAGTCCGATTAAGGAGAGCGTGCAGTATCTATGATTGGAAGTATGGGAATGTT 6286  
QY 5120 GATACCCGATTTCTTCAGTGTCTTCAGGTCTCCTATCAGATTATCCCAACTTAAAGCAAC 5179  
Db 6287 GATACCCGATTTCTTCAGTGTCTTCAGGTCTCCTATCAGATTATCCCAACTTAAAGCAAC 6346  
QY 5180 CGAGAGAGAGATTTCAATGTAATTTCTTCGACTTTTGGTCATCAGTAGACTCGAACTG 5239  
Db 6347 CGAGAGAGAGATTTCAATGTAATTTCTTCGACTTTTGGTCATCAGTAGACTCGAACTG 6406  
QY 5240 TGAGACTATCTCGGTATGACAGCAGAAATGTCCTTCTTGGAGACAGTAATGAAAGTCCC 5299  
Db 6407 TGAGACTATCTCGGTATGACAGCAGAAATGTCCTTCTTGGAGACAGTAATGAAAGTCCC 6466  
QY 5300 ACCAATAAAGAAATCCTTGTATCAGGAACAAACTTCTTGTTCGAACTTTTCGGTCCC 5359  
Db 6467 ACCAATAAAGAAATCCTTGTATCAGGAACAAACTTCTTGTTCGAACTTTTCGGTCCC 6526  
QY 5360 TTGAATATAAATGATAGTGGATATGTCGGTAGGAATCGAGCGGCAATGCTTACC 5419  
Db 6527 TTGAATATAAATGATAGTGGATATGTCGGTAGGAATCGAGCGGCAATGCTTACC 6586  
QY 5420 TTCTGACCTTCAAGAGGTATGAGGTTTGTAGATACCTGATGCCAACTTCAGTGACAAC 5479  
Db 6587 TTCTGACCTTCAAGAGGTATGAGGTTTGTAGATACCTGATGCCAACTTCAGTGACAAC 6646  
QY 5480 GTTGCTATTTCTTCAAAACCATTCGAAATCCAGAGAAATCAAGTTGTTTGTCTACTATT 5539  
Db 6647 GTTGCTATTTCTTCAAAACCATTCGAAATCCAGAGAAATCAAGTTGTTTGTCTACTATT 6706  
QY 5540 GATCAAGCAGTGCCTTCTGAAACTGACAATAGTGTCTGCTGTTTGTAGGTCATCTT 5599  
Db 6707 GATCAAGCAGTGCCTTCTGAAACTGACAATAGTGTGCTGCTGTTTGTAGGTCATCTT 6766  
QY 5600 TGTATGAATAAATCTAGTCTTTGATCTATAATTAATCTCAAGAGCAAGGCAATAATACC 5659  
Db 6767 TGTATGAATAAATCTAGTCTTTGATCTATAATTAATCTCAAGAGCAAGGCAATAATACC 6826  
QY 5660 CAATCTAATAACTCTTTTAAACGTTTAAAGCAGTAAGTATGTCCTCTGTTTAAACCCC 5719  
Db 6827 CAATCTAATAACTCTTTTAAACGTTTAAAGCAGTAAGTATGTCCTCTGTTTAAACCCC 6886  
QY 5720 AAATCAGCTCTGATCTCATCACTTCACTTGAAGGCGCACTATCTTGTTTTAGAGAAAT 5779  
Db 6887 AAATCAGCTCTGATCTCATCACTTCACTTGAAGGCGCACTATCTTGTTTTAGAGAAAT 6946  
QY 5780 TTGCGGAGATGGATATCGAGAAAGATACGCTGATTTTAAACGTAATTTATCTCAA 5839  
Db 6947 TTGCGGAGATGGATATCGAGAAAGATACGCTGATTTTAAACGTAATTTATCTCAA 7006  
QY 5840 GATCTCTGCTCGCGCTTTCGGTGATCAGCGTGAACCTCTGACACATCAGCTCCCG 5899  
Db 7007 GATCTCTGCTCGCGCTTTCGGTGATCAGCGTGAACCTCTGACACATCAGCTCCCG 7066  
QY 5900 GAGACGCTCAGCTGTGCTGTAAGCGATGCGGGAGACAGACGCCCTCAGGGCCCG 5959  
Db 7067 GAGACGCTCAGCTGTGCTGTAAGCGATGCGGGAGACAGACGCCCTCAGGGCCCG 7126  
QY 5960 TCAGCGGTTTGGCGGCTGTCGGGCGCAGCCATGACCCAGTACGTAGCGATAGCGGA 6019  
Db 7127 TCAGCGGTTTGGCGGCTGTCGGGCGCAGCCATGACCCAGTACGTAGCGATAGCGGA 7186  
QY 6020 GTGTATCTGCTTAACTATCGGCTCAGAGCAGATTTGTAAGAGTGCACCATATGC 6079  
Db 7187 GTGTATCTGCTTAACTATCGGCTCAGAGCAGATTTGTAAGAGTGCACCATATGC 7246  
QY 6080 GGTGTGAATACCGCAGATCGGTAAAGGAAATACCGCATCAGGCGCTTCTCCGCTT 6139  
Db 7247 GGTGTGAATACCGCAGATCGGTAAAGGAAATACCGCATCAGGCGCTTCTCCGCTT 7306  
QY 6140 CCTCGCTCACTGACTCGCTGCGCTCGGTGCTGCGCTGCGGCGAGCGGTATCAGCTCACT 6199



LOCUS AX149597 9276 bp DNA linear PAT 31-AUG-2001  
DEFINITION Sequence 21 from Patent WO0136604.  
ACCESSION AX149597  
VERSION AX149597.1 GI:14348030  
KEYWORDS Pichia pastoris  
SOURCE Pichia pastoris  
ORGANISM Pichia pastoris  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Pichia.  
REFERENCE 1  
AUTHORS Madison, E.L. and Ong, E.O.  
TITLE Nucleic acids encoding endotheliases, endotheliases and uses  
thereof  
JOURNAL Patent: WO 0136604-A 21 25-MAY-2001;  
CORVAS INTERNATIONAL, INC. (US)  
FEATURES  
source Location/Qualifiers  
1 . 9276  
/organism="Pichia pastoris"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4922"  
/note="pPic9K vector"  
ORIGIN  
Query Match 79.6%; Score 6561.6; DB 6; Length 9276;  
Best local Similarity 84.7%; Pred. No. 0;  
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;  
QY 1 AGATCTACATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60  
DB 1 AGATCTACATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60  
QY 61 GTCCATTCTCACATAAGTGCCTAAAGCGCAACAGAGGGGATACACTAGCAGACAGCGT 120  
DB 61 GTCCATTCTCACATAAGTGCCTAAAGCGCAACAGAGGGGATACACTAGCAGACAGCGT 120  
QY 121 TGCAAACGAGACCTCAGCTCTCTCTCTCTCAACACCGACATTTTGGCATCCGAAAC 180  
DB 121 TGCAAACGAGACCTCAGCTCTCTCTCTCTCAACACCGACATTTTGGCATCCGAAAC 180  
QY 181 AGCCAGTATTGGCTTGATGAGCTCGCTCATTCGAACTTCCTTATTAGGCTACTA 240  
DB 181 AGCCAGTATTGGCTTGATGAGCTCGCTCATTCGAACTTCCTTATTAGGCTACTA 240  
QY 241 ACACCATGACTTTATTAGCTTGTCTATCTGGCCCCCTGGCGAGGTTTCATGTTGTTA 300  
DB 241 ACACCATGACTTTATTAGCTTGTCTATCTGGCCCCCTGGCGAGGTTTCATGTTGTTA 300  
QY 301 TTTCGGAATGCAACAGCTCCGATTTACCCGAAACATCTCCAGATGAGGCTTTCTG 360  
DB 301 TTTCGGAATGCAACAGCTCCGATTTACCCGAAACATCTCCAGATGAGGCTTTCTG 360  
QY 361 AGTGTGGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTTAAACGCT 420  
DB 361 AGTGTGGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTTAAACGCT 420  
QY 421 GTCTTGGAACTTAATGCAAAAGCGTGATCTCATCCAGATGAATAGTTGGTTG 480  
DB 421 GTCTTGGAACTTAATGCAAAAGCGTGATCTCATCCAGATGAATAGTTGGTTG 480  
QY 481 TTGAAATGCTAAACGCCAGTTGGTCAAAAGAACTTCCAAAGTCCGCAATACCGTTTGT 540  
DB 481 TTGAAATGCTAAACGCCAGTTGGTCAAAAGAACTTCCAAAGTCCGCAATACCGTTTGT 540  
QY 541 CTGTGTTGGTATTAATGCAAGATGCTCAAAATAATCTCATTAATGCTTAGCGCAGTCT 600  
DB 541 CTGTGTTGGTATTAATGCAAGATGCTCAAAATAATCTCATTAATGCTTAGCGCAGTCT 600  
QY 601 CTCATATCGCTTCTGAACCCCGGTGACCTGTGCGGAAACGCAATGGGAAACACCCGCT 660  
DB 601 CTCATATCGCTTCTGAACCCCGGTGACCTGTGCGGAAACGCAATGGGAAACACCCGCT 660  
QY 661 TTTTGGATGATTATGCAATGCTCCACATTTGATGCTTCCAAAGATTCCTGGTGGAAATACT 720  
DB 661 TTTTGGATGATTATGCAATGCTCCACATTTGATGCTTCCAAAGATTCCTGGTGGAAATACT 720  
QY 721 GCTGATAGCTTAACGCTTCATGATCAAAATTTAACTGTCTTAACCCCTACTTGACAGCAAT 780  
DB 721 GCTGATAGCTTAACGCTTCATGATCAAAATTTAACTGTCTTAACCCCTACTTGACAGCAAT 780  
QY 781 ATATAACAGAGGAGGAGCTGCCCTGTCTTAACCCCTTTTATCATCATATTAGCTT 840  
DB 781 ATATAACAGAGGAGGAGCTGCCCTGTCTTAACCCCTTTTATCATCATATTAGCTT 840  
QY 841 ACTTTTCATAATTCGACCTGGTTCCAAATGACAAGCTTTTGAATTTAAGCACTTTAACA 900  
DB 841 ACTTTTCATAATTCGACCTGGTTCCAAATGACAAGCTTTTGAATTTAAGCACTTTAACA 900  
QY 901 CCACTTGAGAGATCAAAACCACTAATTTTGAAGGATCCAAAGATGAGATTTTCT 960  
DB 901 CCACTTGAGAGATCAAAACCACTAATTTTGAAGGATCCAAAGATGAGATTTTCT 960  
QY 961 TCAATTTTACTGCAAGTTTATTCGACATCTCCGCAATGCTGCCAGTCAACACT 1020  
DB 961 TCAATTTTACTGCAAGTTTATTCGACATCTCCGCAATGCTGCCAGTCAACACT 1020  
QY 1021 ACAACAGAGATGAACGCGCAAAATTCGGCTGAAGCTGTCATCGGTACTCAGATT 1080  
DB 1021 ACAACAGAGATGAACGCGCAAAATTCGGCTGAAGCTGTCATCGGTACTCAGATT 1080  
QY 1081 GAAAGGGGATTCGATGTTGCTGTTTGGCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140  
DB 1081 GAAAGGGGATTCGATGTTGCTGTTTGGCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140  
QY 1141 TTTTAAATATCTATTTGCGAGCATCTGCTCTAAAGAGGGGTATCTCTCGAGAA 1200  
DB 1141 TTTTAAATATCTATTTGCGAGCATCTGCTCTAAAGAGGGGTATCTCTCGAGAA 1200  
QY 1201 AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGATCATG 1260  
DB 1201 AGAGAGGCTGAAGCTTA----- 1217  
QY 1261 GCCCCCCGGGAGGAGATAAATTTGTGGCTTCCCGGTGTCACGCCAGCAGCTGACGGAG 1320  
DB 1218 ----- 1217  
QY 1321 AGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTCGCGTGTCTCCACCCCATGSCC 1380  
DB 1218 ----- 1217  
QY 1381 ATCGAGAACTCAAGAAAGAAATGTCCTTCTAACTAGTGGCGTAGAATTCCTAGGG 1440  
DB 1218 -----CGTAGAATTCCTAGGG 1234  
QY 1441 CGGCCGCAATTAATTCGCTTAGCATGACTGTTCTCAGTTCAGTTGGGCACTTACG 1500  
DB 1235 CGGCCGCAATTAATTCGCTTAGCATGACTGTTCTCAGTTCAGTTGGGCACTTACG 1294  
QY 1501 AGAAGACCGCTCTGCTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTGCCCTGAGAG 1560  
DB 1295 AGAAGACCGCTCTGCTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTGCCCTGAGAG 1354  
QY 1561 ATGAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTT 1620  
DB 1355 ATGAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTT 1414  
QY 1621 GTCAATTTGTTTCTTCTCGTAGAGCTTGTCTCCTGATCAGCCTATCTCGAGCTGATGAA 1680  
DB 1415 GTCAATTTGTTTCTTCTCGTAGAGCTTGTCTCCTGATCAGCCTATCTCGAGCTGATGAA 1474  
QY 1681 TATCTTGTGGTAGGGTTTGGGAAATCATTCGAGTTTGTGTTTCTTGGTATTTTCCC 1740  
DB 1475 TATCTTGTGGTAGGGTTTGGGAAATCATTCGAGTTTGTGTTTCTTGGTATTTTCCC 1534  
QY 1741 ACTCCTCTTCAGAGTACAGAGATTTAGTGAGAGTTTGGTTTGTGCAAGCTTATCGATAA 1800  
DB 1535 ACTCCTCTTCAGAGTACAGAGATTTAGTGAGAGTTTGGTTTGTGCAAGCTTATCGATAA 1594

QY	1801	GCTTTAAATGGCGTAGTTTATCACAGTTTAAATTGTCTAAACGAGTCAGGCACCGGTGTATGAA	1860
DB	1595	GCTTTAAATGGGTAGTTTATCACAGTTTAAATTGTCTAAACGAGTCAGGCACCGGTGTATGAA	1654
QY	1861	ATCTAAACAATGGCGTCATCGTCACTCTCGGCACCGTCACCGCTGGATGCTCTGAGGCATAGG	1920
DB	1655	ATCTAAACAATGGCGTCATCGTCACTCTCGGCACCGTCACCGCTGGATGCTCTGAGGCATAGG	1714
QY	1921	CTTGTTTATGCCGTA CTGCGGCGCTCTTGGCGGATATCGTCCATTCGCACAGCATCGC	1980
DB	1715	CTTGTTTATGCCGTA CTGCGGCGCTCTTGGCGGATATCGTCCATTCGCACAGCATCGC	1774
QY	1981	CAGTCATATGGCGTGTCTAGCGCTATATCGTTGATGCAATTTCTATGCCACCCGT	2040
DB	1775	CAGTCATATGGCGTGTCTAGCGCTATATCGTTGATGCAATTTCTATGCCACCCGT	1834
QY	2041	TCTCGGAGCACTGTCGACCGCTTTGGCGCGCCCACTCTGCTCGCTTCTGCTACTCTGG	2100
DB	1835	TCTCGGAGCACTGTCGACCGCTTTGGCGCGCCCACTCTGCTCGCTTCTGCTACTCTGG	1894
QY	2101	AGGCATCTCGACTACCGGATCATGCGGACACACACCGCTCTGCTGGATCTATCGAATCTA	2160
DB	1895	AGGCATCTCGACTACCGGATCATGCGGACACACACCGCTCTGCTGGATCTATCGAATCTA	1954
QY	2161	AATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCOATACGAACCTTAA	2220
DB	1955	AATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCOATACGAACCTTAA	2014
QY	2221	CAGCAATTCGCGTGAGCATCTAGACCTTTCACAGCAGCCAGATCATCACTGCTTGGCCAA	2280
DB	2015	CAGCAATTCGCGTGAGCATCTAGACCTTTCACAGCAGCCAGATCATCACTGCTTGGCCAA	2074
QY	2281	TATGTTTCAGTCCCTCAGGAGTTACGCTCTTGTGTAAGTGATGAATTTCTGGAAGTTGCAG	2340
DB	2075	TATGTTTCAGTCCCTCAGGAGTTACGCTCTTGTGTAAGTGATGAATTTCTGGAAGTTGCAG	2134
QY	2341	TGTTTAACTCCGCTGATGTAGCGGCATATCGTACGTTGCAAGTCTGTTTGGTACCGG	2400
DB	2135	TGTTTAACTCCGCTGATGTAGCGGCATATCGTACGTTGCAAGTCTGTTTGGTACCGG	2194
QY	2401	AGGAGTAATCTCCACAACTCTCTCGGAGTAGGCACCAACAAACACAGATCCAGCGTGT	2460
DB	2195	AGGAGTAATCTCCACAACTCTCTCGGAGTAGGCACCAACAAACACAGATCCAGCGTGT	2254
QY	2461	GTACTTGTATCAACATPAGAAGAGCATCTCGATTTGCAAGATCAAGTGTTCAGGACCGT	2520
DB	2255	GTACTTGTATCAACATPAGAAGAGCATCTCGATTTGCAAGATCAAGTGTTCAGGACCGT	2314
QY	2521	ACTGATTGGAATTTCCAAAGCCCTGCTGATAGTTTGAACCGATAGGTTGTAGAGTGTG	2580
DB	2315	ACTGATTGGAATTTCCAAAGCCCTGCTGATAGTTTGAACCGATAGGTTGTAGAGTGTG	2374
QY	2591	CAATACACTTGGGTACAAATTTCAACCTTTGCGCAACTGCAAGCTTGGTTGTGTAACAGCAT	2640
DB	2375	CAATACACTTGGGTACAAATTTCAACCTTTGCGCAACTGCAAGCTTGGTTGTGTAACAGCAT	2434
QY	2641	CTTCAATTTGCGAAGCTCTTGTCTGTCATATCGACAGCCCAACAGAAATCACCTGGGAAT	2700
DB	2435	CTTCAATTTGCGAAGCTCTTGTCTGTCATATCGACAGCCCAACAGAAATCACCTGGGAAT	2494
QY	2701	CAATACCAATGTTACGCTTGAGACAGAGGTCTGAGGCACGAAATCTTGGATCAGCGTATT	2760
DB	2495	CAATACCAATGTTACGCTTGAGACAGAGGTCTGAGGCACGAAATCTTGGATCAGCGTATT	2554
QY	2761	TATCAGCAATTAACTAGAACTTTCAGAAGGCCACAGAGGCATGTCATATCTACACAGGGCTG	2820
DB	2555	TATCAGCAATTAACTAGAACTTTCAGAAGGCCACAGAGGCATGTCATATCTACACAGGGCTG	2614
QY	2821	ATGTCGTATTTTGAACCAATCTATCTTTGGCAGCAGTAAACGAACTTGTTTCTCGACCAATA	2880
DB	2615	ATGTCGTATTTTGAACCAATCTATCTTTGGCAGCAGTAAACGAACTTGTTTCTCGACCAATA	2674

Qy	2881	TTTTGTGACACTTAGAGAAAGAGTTTCTGTTCCGTAGGCATAGCAGCTACTGCTCTGGGCGC	2941
Db	2875	TTTTTGTACACTTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGCAGCTACTGCTCTGGGCGC	2734
Qy	2941	CTCTGCTAGCAGATACACTTAGCACCAACCTTGTGGGCAACGTTAGATGACCTTCTGGGG	3000
Db	2735	CTCTGCTAGCAGATACACTTTAGCACCACCTTGTGGGCAACGTTAGATGACCTTCTGGGG	2794
Qy	3001	TAAGGFTACCATCTTCTTTAGTGTGAGATGCAAAAAACAATTTCTTTTGCAACGACAACTT	3060
Db	2795	TAAGGFTACCATCTTCTTTAGTGTGAGATGCAAAAAACAATTTCTTTTGCAACGACAACTT	2854
Qy	3061	TGCGAGGAACACCCAGCATCAGGAGAGTGGAAGGCGAGAATTTGGGTTTCCACCCAGGAATAT	3120
Db	2855	TGGCAGGAGACCCAGCATCAGGAGAGTGGAAGGCGAGAATTTGGGTTTCCACCCAGGAATAT	2914
Qy	3121	AGAGGCCAACTTTCTCAATAGGCTTTGCAAAACGAGACGACATACACCCAGGCGCAAGTCT	3180
Db	2915	AGAGGCCAACTTTCTCAATAGGCTTTGCAAAACGAGACGACATACACCCAGGCGCAAGTCT	2974
Qy	3181	CAACTTGCACAGTCTCGTTAGTTGAGCTTCATGGAAATTTCCGTGATCTTATATAGAGA	3240
Db	2975	CAACTTGCACAGTCTCGTTAGTTGAGCTTCATGGAAATTTCCGTGATCTTATATAGAGA	3034
Qy	3241	GATCAATGGCTCTCTTAAACGTTATCTGGCAATTTGCATAAGTTCTCTGGGAAAGGAGCTT	3300
Db	3095	GATCAATGGCTCTCTTAAACGTTATCTGGCAATTTGCATAAGTTCTCTGGGAAAGGAGCTT	3094
Qy	3301	CTAACACAGGTGCTTCAAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT	3360
Db	3095	CTAACACAGGTGCTTCAAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT	3154
Qy	3361	TGTCACCATTTTGACGAAACATTTGCGCAATTTGGTTTGACTAATTCCTAATCTGTTCGG	3420
Db	3155	TGTCACCATTTTGACGAAACATTTGCGCAATTTGGTTTGACTAATTCCTAATCTGTTCGG	3214
Qy	3421	TTTTTCTGGATAGGACGAGGAAGGCATCTTTCAATTTCTTTGTGAGGAGGCGCTTAGAAGCGT	3480
Db	3215	TTTTTCTGGATAGGACGAGGAAGGCATCTTTCAATTTCTTTGTGAGGAGGCGCTTAGAAGCGT	3274
Qy	3481	CAATTTTTCGACATTTCAATACGACCTTCAGAGGAGCTCTTTTAGTTTTCGATTTCTTCTT	3540
Db	3275	CAATTTTTCGACATTTCAATACGACCTTCAGAGGAGCTCTTTTAGTTTTCGATTTCTTCTT	3334
Qy	3541	TAGTTTCTTCTTGGTGTATCTCTGGCTTGGCATCTCTCTTCTTAGTGACCTTTTAGGG	3600
Db	3335	TAGTTTCTTCTTGGTGTATCTCTGGCTTGGCATCTCTCTTCTTAGTGACCTTTTAGGG	3394
Qy	3601	ACTTCAATACAGGTTTCTCTCGACCTCTGCTCCAAAGTGCACACCGTACTTTGGCAGATCTAA	3660
Db	3395	ACTTCAATACAGGTTTCTCTCGACCTCTGCTCCAAAGTGCACACCGTACTTTGGCAGATCTAA	3454
Qy	3661	CTAATGCAAAATAAATAAGTCAGCACATTTCCACGGCTATATCTTCTTGGATTTAGCTT	3720
Db	3455	CTAATGCAAAATAAATAAGTCAGCACATTTCCACGGCTATATCTTCTTGGATTTAGCTT	3514
Qy	3721	CTGCAAGTTCAATGAGTTCTCTCCATAATTTTAGGGTTCAACAAAACCTTCGTCGCAAAATA	3780
Db	3515	CTGCAAGTTCAATGAGTTCTCTCCATAATTTTAGGGTTCAACAAAACCTTCGTCGCAAAATA	3574
Qy	3781	ACCGTTTGGTATAGAACCTTCTGGAGCATTGCTTACGATCCCAAGGTGGCTTCCCA	3840
Db	3575	ACCGTTTGGTATAGAACCTTCTGGAGCATTGCTTACGATCCCAAGGTGGCTTCCCA	3634
Qy	3841	TGGCTCTAAGACCCCTTTGATTGGCCAAAAAGGAAGTGCCTTCCAGTGCACAGAAACCAA	3900
Db	3635	TGGCTCTAAGACCCCTTTGATTGGCCAAAAAGGAAGTGCCTTCCAGTGCACAGAAACCAA	3694
Qy	3901	CACCTGTTGTTGCAACACAAATTTCAAGCAGTCTCCATCACATCCAAATTCGATGCCCA	3960
Db	3695	CACCTGTTGTTGCAACACAAATTTCAAGCAGTCTCCATCACATCCAAATTCGATGCCCA	3754
Qy	3961	GCAACTTTTTCAGTTTGCTCCAGATGTAGCACTTTTATACCAAAACCGTGCAGCAGATT	4020

Db 3755 GCAACTTTTGAGTTGCTCCAGATGACGACCTTTTATACCAAAACCGTGACGAGATT 3814  
Qy 4021 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4080  
Db 3815 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 3874  
Qy 4081 GGCCCAACAGTAATTAGAGAGTGCACCAACAAAGTAGTGAATAGACCAATCGGGGGCT 4140  
Db 3875 GGCCCAACAGTAATTAGAGAGTGCACCAACAAAGTAGTGAATAGACCAATCGGGGGCT 3934  
Qy 4141 CAGTAGTCAAGAGAGCCCAACAAATTTCACTGACAGGGAATTTTGCACATCTTCAGAAA 4200  
Db 3935 CAGTAGTCAAGAGAGCCCAACAAATTTCACTGACAGGGAATTTTGCACATCTTCAGAAA 3994  
Qy 4201 GTTCGTATTTCAGTAGTCAATTTGCCAGCATCAATPAATGGGATPATACCAAGAACAG 4260  
Db 3995 GTTCGTATTTCAGTAGTCAATTTGCCAGCATCAATPAATGGGATPATACCAAGAACAG 4054  
Qy 4261 TGGAGTCACATCTACCAACTTTGGGTCTCAGAAAAGCATAAACAGTTCTACTCCGC 4320  
Db 4055 TGGAGTCACATCTACCAACTTTGGGTCTCAGAAAAGCATAAACAGTTCTACTCCGC 4114  
Qy 4321 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAGGCACAGCGATACTAGCAT 4380  
Db 4115 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAGGCACAGCGATACTAGCAT 4174  
Qy 4381 TAGCGGGAAGATGCAACTTTATCAACAGGGTCTTATAGATAACCTAGCGCTGGGA 4440  
Db 4175 TAGCGGGAAGATGCAACTTTATCAACAGGGTCTTATAGATAACCTAGCGCTGGGA 4234  
Qy 4441 TCATCCTTTGGACAACTTTTCTGCCAAATCTAGTCCAAATCACTTCATTGATACCAT 4500  
Db 4235 TCATCCTTTGGACAACTTTTCTGCCAAATCTAGTCCAAATCACTTCATTGATACCAT 4294  
Qy 4501 TATTGTACAACTTGAGCAAGTTGATCAGCTCCTCAAAATGCTCTGTAAACGATG 4560  
Db 4295 TATTGTACAACTTGAGCAAGTTGATCAGCTCCTCAAAATGCTCTGTAAACGATG 4354  
Qy 4561 ACTCAACTTGACATTAATTTAGACTCAGTCAAGTGGAGAACTTGATCAGGTTGTGCA 4620  
Db 4355 ACTCAACTTGACATTAATTTAGACTCAGTCAAGTGGAGAACTTGATCAGGTTGTGCA 4414  
Qy 4621 GCTGTGACAGCAATGAGGAACACGCGTTTCTCAAACTCAAGAAATTTATCAAACT 4680  
Db 4415 GCTGTGACAGCAATGAGGAACACGCGTTTCTCAAACTCAAGAAATTTATCAAACT 4474  
Qy 4681 CTGCAACACTTGCATGACAGTACGAGGAAATGTCATCTGAACTCGGACGTGAG 4740  
Db 4475 CTGCAACACTTGCATGACAGTACGAGGAAATGTCATCTGAACTCGGACGTGAG 4534  
Qy 4741 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATTTATTCAGTGAATCAGTCAATCAGGAGA 4800  
Db 4535 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATTTATTCAGTGAATCAGTCAATCAGGAGA 4594  
Qy 4801 TCCTCTAGCGGAGCGATCTGGCCGA 4828  
Db 4595 TCCTCTAGCGGAGCGATCTGGCCGA 4654  
Qy 4829 4828  
Db 4655 CTCGTGAAGAGGTGTTGCTGACTCATACAGGCTGAAATCGCCCCCATCATCCAGCCAGA 4714  
Qy 4829 4828  
Db 4715 AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGTGGACCAAGTTGTTGATTTGA 4774  
Qy 4829 4828  
Db 4775 ACTTTTGTTCACAGGAAACGGTCTGGTGTCTGGGAGATGGGTGATCTGATCCTTCA 4834  
Qy 4829 4828

Db 4835 ACTCAGCAAAAGTTCGATTTATTCAACAAAGCCGCGTCCGTCAGTCAAGTCAGCGTAATGCT 4894  
Qy 4829 4828  
Db 4895 CTGCCAGTGTACAAACCAATTAAACCAATTTCTGATTAGAAAAAATCATCGAGCATCAAAATG 4954  
Qy 4829 4828  
Db 4955 AAACTGCAATTTATTCATATCAGGATATCAATACCATAATTTTGA AAAAGCCGTTTCTG 5014  
Qy 4829 4828  
Db 5015 TAATGAGGAGAAAACTCAGGAGCGAGTCCATAGGATGGCAAGATCCTGTATCGGTC 5074  
Qy 4829 4828  
Db 5075 TCGGATTCGAGCTCGTCCAAACATCAATACAACTATTATTTCCCTCGTCAAAAAAATAG 5134  
Qy 4829 4828  
Db 5135 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGCAAAAGCTT 5194  
Qy 4829 4828  
Db 5195 ATGCATTTCTTTCCAGACTTGTTCACAGGCGCAGCCATTACGCTCGTCATCAAAATCACT 5254  
Qy 4829 4828  
Db 5255 CGCATCAACCAACCGTTATTTCATTCGTGATTGCGCTGAGCGAGAGAAATACGCGATC 5314  
Qy 4829 4828  
Db 5315 GCTGTTAAAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGAACAATGCCAG 5374  
Qy 4829 4828  
Db 5375 CGCATCAACATATTTTCACTGANTCAGGATATTTCTTAATACCTGGAATGCTGTTT 5434  
Qy 4829 4828  
Db 5435 CCGGGGATCGCAGTGTGAGTAACCATGCATCATCAGGAGTACGGAATAAAATGCTTGAT 5494  
Qy 4829 4828  
Db 5495 GGTGGAAGAGGCAATAATTCGTCAGCCAGTTTGTAGTCTGACCATCTCATCTGTAAACATC 5554  
Qy 4829 4828  
Db 5555 ATTGCAACGCTACCTTTGCCATGTTTCAGAAAAAACAATCTGGCGCATCGGCTTCCCATATA 5614  
Qy 4829 4828  
Db 5615 CAATCGATAGATTTGTCGACCTGATTGCCGACATTAATCGGAGCCCATTTATACCCATA 5674  
Qy 4829 4828  
Db 5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGGCTCGAGCAAGAGCTTTCCCGTTGAAT 5734  
Qy 4829 4828  
Db 5735 ATGGCTCATACACCCCTTTGATTATTACTGTTTATGTAAGCAGACAGTTTATTTGTTTCATGA 5794  
Qy 4829 4828  
Db 5795 TGATATATTTTATCTTGTGCAATGTAATCATCAGAGATTTTGAACACACACGTCGCTTTC 5854  
Qy 4829 4828  
Db 5855 CCCCCCCCCCTGCAAGTGGCATCACCGGCGCCACAGTGGCTGCTGGCGCTATAT 5914  
Qy 4880 CCGCGACATCACCGATGGGAGATCGGCTCGCCACTTCGGGCTCATGAGCGCTTGT 4939  
Db 5915 CCGCGACATCACCGATGGGAGATCGGCTCGCCACTTCGGGCTCATGAGCGCTTGT 5974

QY 4940 CGCGTGGGTATGTGGCAGGCCCCGCTGGCGGGGAGCTGTGGCGCCATCTCCTTGCA 4999  
Db 5975 CGCGTGGGTATGTGGCAGGCCCCGCTGGCGGGGAGCTGTGGCGCCATCTCCTTGCA 6034  
QY 5000 TGCACATTCCTTGGCGGGGCTGCTCAAAGCGCTCAACCTACTACTAGGCTGCTTCCT 5059  
Db 6035 TGCACATTCCTTGGCGGGGCTGCTCAAAGCGCTCAACCTACTACTAGGCTGCTTCCT 6094  
QY 5060 AATCAGGAGTGCATATAAGGAGAGCGTCGAGTATCTATGATGGAAATGGT 5119  
Db 6095 AATCAGGAGTGCATATAAGGAGAGCGTCGAGTATCTATGATGGAAATGGT 6154  
QY 5120 GATACCCGATCTTCAGTGTCTTCAGGTCTTCCTATCAGATATGCCCAACTAAGCAAC 5179  
Db 6155 GATACCCGATCTTCAGTGTCTTCAGGTCTTCCTATCAGATATGCCCAACTAAGCAAC 6214  
QY 5180 CGGAGGAGATTCATGGTAAATTTCTCTGACTTTTGGTTCATCAGTAGACTCGAATG 5239  
Db 6215 CGGAGGAGATTCATGGTAAATTTCTCTGACTTTTGGTTCATCAGTAGACTCGAATG 6274  
QY 5240 TGAGACTATCTGGTATGACAGCAGAAATGTCTTCTTGAGACAGTAATGAGTCCC 5299  
Db 6275 TGAGACTATCTGGTATGACAGCAGAAATGTCTTCTTGAGACAGTAATGAGTCCC 6334  
QY 5300 ACCAATAAGAAATCCTTGTATCAGGAACAAATCTTCTTGTTCGAACTTTTCGGTGCC 5359  
Db 6335 ACCAATAAGAAATCCTTGTATCAGGAACAAATCTTCTTGTTCGAACTTTTCGGTGCC 6394  
QY 5360 TTGAATCTAAATAGAGTGATATGTCGGGTAGGAATCGAGCGGCAAACTCTTACC 5419  
Db 6395 TTGAATCTAAATAGAGTGATATGTCGGGTAGGAATCGAGCGGCAAACTCTTACC 6454  
QY 5420 TTCTGAGCTTCAAGAGTATGAGGTTTGTAGTACTGATGCCAATTCAGTGACAAC 5479  
Db 6455 TTCTGAGCTTCAAGAGTATGAGGTTTGTAGTACTGATGCCAATTCAGTGACAAC 6514  
QY 5480 GTTGTATTTTGGTTCAGAACCTTCGGAATCCAGAGAAATCAAAGTTGTTTGTCTACTAT 5539  
Db 6515 GTTGTATTTTGGTTCAGAACCTTCGGAATCCAGAGAAATCAAAGTTGTTTGTCTACTAT 6574  
QY 5540 GATCCAGCCAGTGGCTTGAACCTGACAAATAGTGTGCTCGTGTGAGGTCACTT 5599  
Db 6575 GATCCAGCCAGTGGCTTGAACCTGACAAATAGTGTGCTCGTGTGAGGTCACTT 6634  
QY 5600 TGATGAATAAATCTAGTCTTGTATTAATAATCTTCAGCAGCAAGGCGATAAATACC 5659  
Db 6635 TGATGAATAAATCTAGTCTTGTATTAATAATCTTCAGCAGCAAGGCGATAAATACC 6694  
QY 5660 CAAATCTAAATCTTTTAAACCGTTAAAGCAAGATGTCCTGATTAATTAACCCC 5719  
Db 6695 CAAATCTAAATCTTTTAAACCGTTAAAGCAAGATGTCCTGATTAATTAACCCC 6754  
QY 5720 AAATCAGCTCGTAGTCTCATCAATCTTGAAGGCGCACTATCTTGTGAGAAAT 5779  
Db 6755 AAATCAGCTCGTAGTCTCATCAATCTTGAAGGCGCACTATCTTGTGAGAAAT 6814  
QY 5780 TTGCGGAGATCGATATCAGAAAAGGTACGCTGATTTAAAGCGTAAATTTATCTCAA 5839  
Db 6815 TTGCGGAGATCGATATCAGAAAAGGTACGCTGATTTAAAGCGTAAATTTATCTCAA 6874  
QY 5840 GATCTCTGCTCGCGCTTCGGTGATCAGCGTGAAACCTCTGACATGCACTCCG 5899  
Db 6875 GATCTCTGCTCGCGCTTCGGTGATCAGCGTGAAACCTCTGACATGCACTCCG 6934  
QY 5900 GAGACGTCACAGCTTGTCTGTAGCGGATGCGGGAGCAGCAAGCCGTCAGGGCGG 5959  
Db 6935 GAGACGTCACAGCTTGTCTGTAGCGGATGCGGGAGCAGCAAGCCGTCAGGGCGG 6994  
QY 5960 TCAGCGGTGTTGGCGGTGTCGGGCGCAGCCATGACCCAGTCAGTCAGGATAGCGGA 6019  
Db 6995 TCAGCGGTGTTGGCGGTGTCGGGCGCAGCCATGACCCAGTCAGTCAGGATAGCGGA 7054

QY 6020 GTGTATATCTGCCTTAATATGCGGCATCAGAGCAGATTGTACTGAGAGTGCAACATATGC 6079  
Db 7055 GTGTATATCTGCCTTAATATGCGGCATCAGAGCAGATTGTACTGAGAGTGCAACATATGC 7114  
QY 6080 GGTGTGAAATACCGCAGACAGATCGTAAGAGGAAATACCGCATCAGCGCTCTTCGCTT 6139  
Db 7115 GGTGTGAAATACCGCAGACAGATCGTAAGAGGAAATACCGCATCAGCGCTCTTCGCTT 7174  
QY 6140 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCACT 6199  
Db 7175 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCACT 7234  
QY 6200 CAAAGCGGTAAATACGGTTATCCACAGATCAGGGATTAACGAGGAAAGAAATGTGAG 6259  
Db 7235 CAAAGCGGTAAATACGGTTATCCACAGATCAGGGATTAACGAGGAAAGAAATGTGAG 7294  
QY 6260 CAAAAGCCAGCAAAAAGGCGCAGGAAACCGTAAAGGCGCGTTGCTGGGCTTTTCCATA 6319  
Db 7295 CAAAAGCCAGCAAAAAGGCGCAGGAAACCGTAAAGGCGCGTTGCTGGGCTTTTCCATA 7354  
QY 6320 GGCTCGCCCTCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC 6379  
Db 7355 GGCTCGCCCTCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC 7414  
QY 6380 CGACAGACTATATAAGATACAGGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCTCCCTG 6439  
Db 7415 CGACAGACTATATAAGATACAGGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCTCCCTG 7474  
QY 6440 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGAAAGCTGGCGC 6499  
Db 7475 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGAAAGCTGGCGC 7534  
QY 6500 TTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCCAAGCTGG 6559  
Db 7535 TTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCCAAGCTGG 7594  
QY 6560 GCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCTAACTATCGTC 6619  
Db 7595 GCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCTAACTATCGTC 7654  
QY 6620 TTGAGTCCAAACCGGTAAAGACAGCTTATCGCAGCTGCGAGCCACTGTTAAACGGA 6679  
Db 7655 TTGAGTCCAAACCGGTAAAGACAGCTTATCGCAGCTGCGAGCCACTGTTAAACGGA 7714  
QY 6680 TTAGCAGAGCAGGATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAATACG 6739  
Db 7715 TTAGCAGAGCAGGATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAATACG 7774  
QY 6740 GCTACACTAGAAAGCAGTATTTGGTATCTGGCTCTGCTGAAGCCAGTTAOCCTTCGAA 6799  
Db 7775 GCTACACTAGAAAGCAGTATTTGGTATCTGGCTCTGCTGAAGCCAGTTAOCCTTCGAA 7834  
QY 6800 AAAGAGTTGGTAGTCTTGTATCCGCAACAAACCAACCCGCTGCTAGCGGTGTTTTTGG 6859  
Db 7835 AAAGAGTTGGTAGTCTTGTATCCGCAACAAACCAACCCGCTGCTAGCGGTGTTTTTGG 7894  
QY 6860 TTTGCAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAGATCCTTGAATCTTTT 6919  
Db 7895 TTTGCAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAGATCCTTGAATCTTTT 7954  
QY 6920 CTACGGGTCTGACGCTCAGTGGAACGAAACTCAGCTTAAGGATTTTGGTCAAGAT 6979  
Db 7955 CTACGGGTCTGACGCTCAGTGGAACGAAACTCAGCTTAAGGATTTTGGTCAAGAT 8014  
QY 6980 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATCAATCT 7039  
Db 8015 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATCAATCT 8074  
QY 7040 AAAGTATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAATCAGTAGGCACTTA 7099  
Db 8075 AAAGTATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAATCAGTAGGCACTTA 8134  
QY 7100 TCTCAGCGATGCTCTATTTCGTTTCATCATAGTTGTCCTGACTCCCGCTGCTAGATAA 7159



Db 8135 TCTCAGCGATCTCTCTATTTGCTTCATCCATAGTTGCCGTGCTAGATAA 8194  
Qy 7160 CTACGATACGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGGAGACCCAC 7219  
Db 8195 CTACGATACGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGGAGACCCAC 8254  
Qy 7220 GCTCAGCGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGGCGGAGCGAGAA 7279  
Db 8255 GCTCAGCGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGGCGGAGCGAGAA 8314  
Qy 7280 GTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAATTAATTTGCGGGAAGCTAGAG 7339  
Db 8315 GTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAATTAATTTGCGGGAAGCTAGAG 8374  
Qy 7340 TAAGTAGTTCCGCAAGTAAATAGTTTGCSCAAGTTTGTGCCATTCCTCAGGCAATCTGG 7399  
Db 8375 TAAGTAGTTCCGCAAGTAAATAGTTTGCSCAAGTTTGTGCCATTCCTCAGGCAATCTGG 8434  
Qy 7400 TGTCAGCGCTCGTCTGTTGATGGCTTCATTTCAGCTCCGGTTCCCAACGATCAAGGCGAG 7459  
Db 8435 TGTCAGCGCTCGTCTGTTGATGGCTTCATTTCAGCTCCGGTTCCCAACGATCAAGGCGAG 8494  
Qy 7460 TTACATGATCCCCCATGTTGCGCAAAAAGGGTTAGCTCTTCGGTCTCCGATCTGTTG 7519  
Db 8495 TTACATGATCCCCCATGTTGCGCAAAAAGGGTTAGCTCTTCGGTCTCCGATCTGTTG 8554  
Qy 7520 TCAGAAAGTAAGTTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGCATTAATCTC 7579  
Db 8555 TCAGAAAGTAAGTTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGCATTAATCTC 8614  
Qy 7580 TTACTGTATGCCATCCGTAAGATGCTTTTCTGAGCTGGTGTAGTACTCAACCAAGTCAT 7639  
Db 8615 TTACTGTATGCCATCCGTAAGATGCTTTTCTGAGCTGGTGTAGTACTCAACCAAGTCAT 8674  
Qy 7640 TCTGAGAATAGTGTATCGCGGACGAGTGTCTCTTGGCCCGGCTCAACACGGGATAATA 7699  
Db 8675 TCTGAGAATAGTGTATCGCGGACGAGTGTCTCTTGGCCCGGCTCAACACGGGATAATA 8734  
Qy 7700 CCGCGCCATACGAGAACTTTAAAGTGTCTCATCTTGGAAAGCTTCTTCGGGGGAA 7759  
Db 8735 CCGCGCCATACGAGAACTTTAAAGTGTCTCATCTTGGAAAGCTTCTTCGGGGGAA 8794  
Qy 7760 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTCAACCCA 7819  
Db 8795 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTCAACCCA 8854  
Qy 7820 ACTGATCTTCAGCATCTTTTACTTTTCAGAGGTTTCTGGGTGAGCAAAAACAGGAAGGC 7879  
Db 8855 ACTGATCTTCAGCATCTTTTACTTTTCAGAGGTTTCTGGGTGAGCAAAAACAGGAAGGC 8914  
Qy 7880 AAAATGCCCAAAAAGGGAATAAGGGGACACGGAATGTTGAATACTCATCTCTTCC 7939  
Db 8915 AAAATGCCCAAAAAGGGAATAAGGGGACACGGAATGTTGAATACTCATCTCTTCC 8974  
Qy 7940 TTTTTCATATTTATGAGCAATTTATCAGGTTTATGCTCATGAGCGGATACATATTG 7999  
Db 8975 TTTTTCATATTTATGAGCAATTTATCAGGTTTATGCTCATGAGCGGATACATATTG 9034  
Qy 8000 AATGTATTTAGAAAATAAACAATAGGGGTTCCCGGCACATTTCCCGCAAAAGTGCAC 8059  
Db 9035 AATGTATTTAGAAAATAAACAATAGGGGTTCCCGGCACATTTCCCGCAAAAGTGCAC 9094  
Qy 8060 CTGAGCTCTAAGAACCAATTTATCATGACATTAACCTATAAATAAGCGGTATCAGGA 8119  
Db 9095 CTGAGCTCTAAGAACCAATTTATCATGACATTAACCTATAAATAAGCGGTATCAGGA 9154  
Qy 8120 GGCCTCTTCGTTCTTCAAGAAATTAATCTCATGTTTGCAGCTTATCATGTAAGCTGAC 8179  
Db 9155 GGCCTCTTCGTTCTTCAAGAAATTAATCTCATGTTTGCAGCTTATCATGTAAGCTGAC 9214  
Qy 8180 TCATGTTGGTATGTGCAATAGACCGCATCGGAACTGGAATAAATAACAGTTATTT 8239

Db 9215 TCATGTTGGTATGTGAAATAGACGAGATCGGGAACACTGAAATAAACAGTTATTATT 9274  
Qy 8240 CG 8241  
Db 9275 CG 9276  
RESULT 8  
AX207941  
LOCUS 9276 bp DNA linear PAT 31-AUG-2001  
DEFINITION Sequence 45 from Patent WO0157194.  
ACCESSION AX207941  
VERSION AX207941.1 GI:15422539  
KEYWORDS Pichia pastoris  
SOURCE Pichia pastoris  
ORGANISM Pichia pastoris  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Pichia.  
REFERENCE 1  
AUTHORS Madison, E.L., Ong, E.O. and Yeh, J.C.  
TITLE Nucleic acid molecules encoding transmembrane serine proteases, the  
encoded proteins and methods based thereon  
JOURNAL Patent: WO 0157194-A 45 09-AUG-2001;  
CORVAS INTERNATIONAL, INC. (US)  
FEATURES  
Location/Qualifiers  
1. 9276  
/organism="Pichia pastoris"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4922"  
ORIGIN  
Query Match 79.6%; Score 6561.6; DB 6; Length 9276;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;  
Qy 1 AGATCTAATCAATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAG 60  
Db 1 AGATCTAATCAATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAG 60  
Qy 61 GTCCATTTCTCACATTAAGTGCACAAACGAAACGAGGAGGATACACTAGCAGACGCGT 120  
Db 61 GTCCATTTCTCACATTAAGTGCACAAACGAAACGAGGAGGATACACTAGCAGACGCGT 120  
Qy 121 TGCAAAACGAGGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGGCATCGAAAACC 180  
Db 121 TGCAAAACGAGGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGGCATCGAAAACC 180  
Qy 181 AGCCGAGTTATTGGGCTTGATTTGAGCTCGCTCATTCGAAATTCCTTCTATTAGGCTACTA 240  
Db 181 AGCCGAGTTATTGGGCTTGATTTGAGCTCGCTCATTCGAAATTCCTTCTATTAGGCTACTA 240  
Qy 241 ACACCATGACTTTATTAGCCTGTCTATCTCGGCCCTCGGCGAGGTTTCATGTTGTTTA 300  
Db 241 ACACCATGACTTTATTAGCCTGTCTATCTCGGCCCTCGGCGAGGTTTCATGTTGTTTA 300  
Qy 301 TTTCCGAATGCAACAGCTCCGATTTACCCCGAATACATCCAGATGAGGCTTTCTG 360  
Db 301 TTTCCGAATGCAACAGCTCCGATTTACCCCGAATACATCCAGATGAGGCTTTCTG 360  
Qy 361 AGTGTGGGTTCAATAGTTTTCATGTTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420  
Db 361 AGTGTGGGTTCAATAGTTTTCATGTTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420  
Qy 421 GTCTTGAACTTAATATGACAAAGCGGTGATCTCATCCAGATGAATGAATTTGGTTCG 480  
Db 421 GTCTTGAACTTAATATGACAAAGCGGTGATCTCATCCAGATGAATGAATTTGGTTCG 480  
Qy 481 TTGAATGCTTAACGGCGAGTTGGTCAAAAAGAACTTCCAAAAGTCGCGCATACCGTTGT 540  
Db 481 TTGAATGCTTAACGGCGAGTTGGTCAAAAAGAACTTCCAAAAGTCGCGCATACCGTTGT 540  
Qy 541 CTCTGTTGGTATTGATGAGGAATGCTCAAAAATAATCTCATTAATGCTTAGCGAGTCT 600

Db 541 CTTGTTTGGTATGATTGACGAATGCTCAAAATAAATCTCAATTAATGCTTAGCGAGTCT 600  
Qy 601 CTCATCGCTTCTGAACCCCGTGCACCTGTGCGGAAACGCAAAATGGGAAACACCCGCT 660  
Db 601 CTCATCGCTTCTGAACCCCGTGCACCTGTGCGGAAACGCAAAATGGGAAACACCCGCT 660  
Qy 661 TTTTGGATGATTATGATGCTCTCACATTTGATGCTTCCAGATTCTGGTGGGAATACT 720  
Db 661 TTTTGGATGATTATGATGCTCTCACATTTGATGCTTCCAGATTCTGGTGGGAATACT 720  
Qy 721 GCTGATAGCCTAAACCTTCATGATCAAAATTAACCTGTCTAACCCCTACTTGCAGCAAT 780  
Db 721 GCTGATAGCCTAAACCTTCATGATCAAAATTAACCTGTCTAACCCCTACTTGCAGCAAT 780  
Qy 781 ATATAAACAAGAGAGCTGCCCTGTCTTAACCTTTTTTTTATATCATATTAAGCTT 840  
Db 781 ATATAAACAAGAGAGCTGCCCTGTCTTAACCTTTTTTTTATATCATATTAAGCTT 840  
Qy 841 ACTTTCAATATTTGCCACTTGTCTCAATTAACAAGCTTTTGAATTTAAACGACTTTTAACGA 900  
Db 841 ACTTTCAATATTTGCCACTTGTCTCAATTAACAAGCTTTTGAATTTAAACGACTTTTAACGA 900  
Qy 901 CAACCTTGAGAGATCAAAACCAACTAATTAATTTGAAGAGTCCAAAGATGAGATTTCCT 960  
Db 901 CAACCTTGAGAGATCAAAACCAACTAATTAATTTGAAGAGTCCAAAGATGAGATTTCCT 960  
Qy 961 TCAATTTTACTGCACTTTTATTTGCGAGCATCTCCGCAATGCTGCTCCAGTCAACACT 1020  
Db 961 TCAATTTTACTGCACTTTTATTTGCGAGCATCTCCGCAATGCTGCTCCAGTCAACACT 1020  
Qy 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTATCGGTTACTCAGATTTA 1080  
Db 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTATCGGTTACTCAGATTTA 1080  
Qy 1081 GAAGGGATTTGATGCTGTTGCTGTTTGGCATTTTCCACAGACAAATAACGGGTTATTG 1140  
Db 1081 GAAGGGATTTGATGCTGTTGCTGTTTGGCATTTTCCACAGACAAATAACGGGTTATTG 1140  
Qy 1141 TTTATAAATACTACTATTTCCAGCATTTCTGCTAAGAGAGAGGGGTATCTCTCGAGAA 1200  
Db 1141 TTTATAAATACTACTATTTCCAGCATTTCTGCTAAGAGAGAGGGGTATCTCTCGAGAA 1200  
Qy 1201 AGAGAGCTGAAGCCCGAGCCCGAGCCCGAGCCCGAGGAGAGAAATGATCATG 1260  
Db 1201 AGAGAGCTGAAGCTTA ----- 1217  
Qy 1261 GCGCCCGGAGAGAGATAAATTTGGCTTCCCGGTGTACCGCCGAGCTGCAACGGAG 1320  
Db 1218 ----- 1217  
Qy 1321 AGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTTCCGTTGGTGTCTCCACCCATGGCC 1380  
Db 1218 ----- 1217  
Qy 1381 ATCGAGAACCTCAAGAGAGAGATGTCCTTCTAACTAGTGGCGTAGAATTCCTAGGG 1440  
Db 1218 -----CGTAGAATTCCTAGGG 1234  
Qy 1441 CGGCGCGCAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTTGGGCACTTACG 1500  
Db 1235 CGGCGCGCAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTTGGGCACTTACG 1294  
Qy 1501 AGAAGACCGCTTCTGCTAGATTCTAATCAAGAGAGATGTCAGAAATTCCTGAGAG 1560  
Db 1295 AGAAGACCGCTTCTGCTAGATTCTAATCAAGAGAGATGTCAGAAATTCCTGAGAG 1354  
Qy 1561 ATGCAAGCTTCAATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1620  
Db 1355 ATGCAAGCTTCAATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1414  
Qy 1621 GTCAATTTGTTCTCTGAGAGCTTCTCTGATCAGCTATCTCGAGCTGATGAA 1680  
Db 1415 GTCAATTTGTTCTCTGAGAGCTTCTCTGATCAGCTATCTCGAGCTGATGAA 1474

Qy 1681 TATCTTTGCTAGGGTTTGGGAAATCAATCGAGTTTGATGTTTCTTGTTATTTCCC 1740  
Db 1475 TATCTTTGCTAGGGTTTGGGAAATCAATCGAGTTTGATGTTTCTTGTTATTTCCC 1534  
Qy 1741 ACTCTCTTTCAGAGTACAGAGATTAAAGTACAGAGTTCTGTTGTCAGAGCTTATCCATAA 1800  
Db 1535 ACTCTCTTTCAGAGTACAGAGATTAAAGTACAGAGTTCTGTTGTCAGAGCTTATCCATAA 1594  
Qy 1801 GCTTTAATGCGTATGTTTATACAGTTAAATTTGCTAAACAGTACAGCACCGTGTATGAA 1860  
Db 1595 GCTTTAATGCGTATGTTTATACAGTTAAATTTGCTAAACAGTACAGCACCGTGTATGAA 1654  
Qy 1861 ATCTAAATGCGCTCATCGTCACTCTCGCACCGCTCACCTGGATGCTGTAGGATAGG 1920  
Db 1655 ATCTAAATGCGCTCATCGTCACTCTCGCACCGCTCACCTGGATGCTGTAGGATAGG 1714  
Qy 1921 CTTGGTTATGCGCGGTACTGCGCGGCTCTTTGGGGATATCGTCCATTCGCAAGCATCGC 1980  
Db 1715 CTTGGTTATGCGCGGTACTGCGCGGCTCTTTGGGGATATCGTCCATTCGCAAGCATCGC 1774  
Qy 1981 CAGTCACTATGCGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGT 2040  
Db 1775 CAGTCACTATGCGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGT 1834  
Qy 2041 TCTCGAGACACTGCTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTTCTGCTACTTGG 2100  
Db 1835 TCTCGAGACACTGCTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTTCTGCTACTTGG 1894  
Qy 2101 AGCCACTATGACTACCGGATCATGCGCACCAACACCGCTCTGTGGATCTATCGAATCTA 2160  
Db 1895 AGCCACTATGACTACCGGATCATGCGCACCAACACCGCTCTGTGGATCTATCGAATCTA 1954  
Qy 2161 AATGTAAGTTAAATCTCTAAATAAATAAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2220  
Db 1955 AATGTAAGTTAAATCTCTAAATAAATAAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2014  
Qy 2221 CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGCCCAA 2280  
Db 2015 CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGCCCAA 2074  
Qy 2281 TATGTTTCTAGTCCCTCAGGAGTTACGCTCTTGTGAAGTATGAACCTTCTGGAAGTTGCGAG 2340  
Db 2075 TATGTTTCTAGTCCCTCAGGAGTTACGCTCTTGTGAAGTATGAACCTTCTGGAAGTTGCGAG 2134  
Qy 2341 TGTAACTCCGCTGATTTGACGGGCATATCCGTAAGTTGGCAAGTGTGGTTGGTACCGG 2400  
Db 2135 TGTAACTCCGCTGATTTGACGGGCATATCCGTAAGTTGGCAAGTGTGGTTGGTACCGG 2194  
Qy 2401 AGAGTAAATCTCCAACTCTCTGAGAGTATGACCAACAAACAGATCCAGCGTGT 2460  
Db 2195 AGAGTAAATCTCCAACTCTCTGAGAGTATGACCAACAAACAGATCCAGCGTGT 2254  
Qy 2461 GTACTTTGATCAACATGAAGAGAGCATTCGATTTGCAAGTCAAGTCTTCAGAGCGCT 2520  
Db 2255 GTACTTTGATCAACATGAAGAGAGCATTCGATTTGCAAGTCAAGTCTTCAGAGCGCT 2314  
Qy 2521 ACTGATTGAGCAATTTCCAAAGCCTGCTGATGTTGCAACCCGATAGAGTGTG 2580  
Db 2315 ACTGATTGAGCAATTTCCAAAGCCTGCTGATGTTGCAACCCGATAGAGTGTG 2374  
Qy 2581 CAATACACTTGGTACAAATTTCAACCTTGGCACTGCAAGCTTGGTTGTGAACAGCAT 2640  
Db 2375 CAATACACTTGGTACAAATTTCAACCTTGGCACTGCAAGCTTGGTTGTGAACAGCAT 2434  
Qy 2641 CTTCAATTTGCGAAGCTCTTGTCTGTCATATCGACAGCCCAACAGAAATCACTGGGAAT 2700  
Db 2435 CTTCAATTTGCGAAGCTCTTGTCTGTCATATCGACAGCCCAACAGAAATCACTGGGAAT 2494  
Qy 2701 CAATACACTTGGTACAGCTTGAAGAGTCTGAGGCAACGAAATCTGGATCAGCGTAT 2760  
Db 2495 CAATACACTTGGTACAGCTTGAAGAGTCTGAGGCAACGAAATCTGGATCAGCGTAT 2554

QY 2761 TATCAGCAATAACTAGAACTTCAGAGGCCAGCAGGCATGTCAATACATACACAGGGCTG 2820  
 Db 2555 TATCAGCAATAACTAGAACTTCAGAGGCCAGCAGGCATGTCAATACATACACAGGGCTG 2614  
 QY 2821 ATGTGTCAATTTGAACCAATCATCTTGGCAGCAATAGCAAGCTGGTTCTTGGACCAATA 2880  
 Db 2615 ATGTGTCAATTTGAACCAATCATCTTGGCAGCAATAGCAAGCTGGTTCTTGGACCAATA 2674  
 QY 2881 TTTTGTCACTTAGGAACAGTTCTCTTCCGTAAGCCATAGCAGCTACTGCTCGGGCG 2940  
 Db 2675 TTTTGTCACTTAGGAACAGTTCTCTTCCGTAAGCCATAGCAGCTACTGCTCGGGCG 2734  
 QY 2941 CTCCTGTAGCAGCATACATAGCAACCACTTGTGGGCAAGTATAGTACTTCTGGGG 3000  
 Db 2735 CTCCTGTAGCAGCATACATAGCAACCACTTGTGGGCAAGTATAGTACTTCTGGGG 2794  
 QY 3001 TAAGGTACCACTCTTCTTAGTGGAGATGCAAAACAAATTTCTTGCACACAGCAACT 3060  
 Db 2795 TAAGGTACCACTCTTCTTAGTGGAGATGCAAAACAAATTTCTTGCACACAGCAACT 2854  
 QY 3061 TGGCAGGAACACCCAGCATCAGGGAAGTGAAGGAGCAATTTGGGGTTCCACAGGAATAT 3120  
 Db 2855 TGGCAGGAACACCCAGCATCAGGGAAGTGAAGGAGCAATTTGGGGTTCCACAGGAATAT 2914  
 QY 3121 AGAGGCCAATTTCTCAATAGTCTTGCAGCAACGAGAGCAGCTACACAGGCGAGTCT 3180  
 Db 2915 AGAGGCCAATTTCTCAATAGTCTTGCAGCAACGAGAGCAGCTACACAGGCGAGTCT 2974  
 QY 3181 CAACCTGCAAGCTCTCGTGTAGTGGATTCATGGAAATTTCTGACGTTATCTATAGAGA 3240  
 Db 2975 CAACCTGCAAGCTCTCGTGTAGTGGATTCATGGAAATTTCTGACGTTATCTATAGAGA 3034  
 QY 3241 GATCAATGGCTCTCTTAACGTTATCTGCAATTTGCATTAAGTTCCTCTGGGAAGGAGCT 3300  
 Db 3035 GATCAATGGCTCTCTTAACGTTATCTGCAATTTGCATTAAGTTCCTCTGGGAAGGAGCT 3094  
 QY 3301 CTACACAGGCTCTCTTAACGCTATCCATCAACCTTGGCAGTTAGTTCTTAAGAGGCT 3360  
 Db 3095 CTACACAGGCTCTCTTAACGCTATCCATCAACCTTGGCAGTTAGTTCTTAAGAGGCT 3154  
 QY 3361 TGTCAACATTTTGACGAACATTTGCGCAATTTGCTTTGACTTAATTCATTAATCTGTTCG 3420  
 Db 3155 TGTCAACATTTTGACGAACATTTGCGCAATTTGCTTTGACTTAATTCATTAATCTGTTCG 3214  
 QY 3421 TTTTCTGATAGGAGGAGGAGGATCTCAATTTCTTGTGAGGAGGCTTGAACAGT 3480  
 Db 3215 TTTTCTGATAGGAGGAGGAGGATCTCAATTTCTTGTGAGGAGGCTTGAACAGT 3274  
 QY 3481 CAATTTTGCACAAATCAATAGCACTTCAGAGGAGCTCTTTAGGTTTGGATTTCTTCT 3540  
 Db 3275 CAATTTTGCACAAATCAATAGCACTTCAGAGGAGCTCTTTAGGTTTGGATTTCTTCT 3334  
 QY 3541 TAGGTTTCTTGGTGTATCTGCTGGCTGGATCTCTTCTTCTTCTTCTTCTTCTTCTTCT 3600  
 Db 3335 TAGGTTTCTTGGTGTATCTGCTGGCTGGATCTCTTCTTCTTCTTCTTCTTCTTCTTCT 3394  
 QY 3601 ACTTCATATCCAGGTTTCTCTCCACTCTGCTCAAGCTCAGACCGTACTTGGCAGATCTAA 3660  
 Db 3395 ACTTCATATCCAGGTTTCTCTCCACTCTGCTCAAGCTCAGACCGTACTTGGCAGATCTAA 3454  
 QY 3661 CTAATGCAAAATAAATAAGTACAGCAATCCAGGCTATATCTTCTTGGATTTAGCTT 3720  
 Db 3455 CTAATGCAAAATAAATAAGTACAGCAATCCAGGCTATATCTTCTTGGATTTAGCTT 3514  
 QY 3721 CTGCAAGTTCATCAGCTTCTCTCTTAATTTTAGGTTTCAACAAACTTCTGCTGCAATA 3780  
 Db 3515 CTGCAAGTTCATCAGCTTCTCTCTCTTAATTTTAGGTTTCAACAAACTTCTGCTGCAATA 3574  
 QY 3781 ACCGTTTGGTATAGAACTTCTGAGCAATTTCTTACGATCCACAGGTTGCTTCCA 3840  
 Db 3575 ACCGTTTGGTATAGAACTTCTGAGCAATTTCTTACGATCCACAGGTTGCTTCCA 3634  
 QY 3841 TGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAGTGGTTCCAAAGTGACAGAAACCA 3900

Db 3635 TGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAAGTGCCTTCCAGTGCAGAAACCA 3694  
 QY 3901 CACTGTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3960  
 Db 3695 CACTGTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3754  
 QY 3961 GCAACTTTTGTAGTTGCTCCAGATGTAGCACTTTATACCAAAACCCGTGACGAGAT 4020  
 Db 3755 GCAACTTTTGTAGTTGCTCCAGATGTAGCACTTTATACCAAAACCCGTGACGAGAT 3814  
 QY 4021 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGAGAGTACCA 4080  
 Db 3815 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGAGAGTACCA 3874  
 QY 4081 GGGCCAAACGAGTAATTAAGAGTCAAGCAGCAACCAAGTAGTGAATAGACCATCGGGCGGT 4140  
 Db 3875 GGGCCAAACGAGTAATTAAGAGTCAAGCAGCAACCAAGTAGTGAATAGACCATCGGGCGGT 3934  
 QY 4141 CAGTAGTCAAGAGCGCCAAACAAATTTCACTGACAGGAACTTTTTCACATCTTCAAAA 4200  
 Db 3935 CAGTAGTCAAGAGCGCCAAACAAATTTCACTGACAGGAACTTTTTCACATCTTCAAAA 3994  
 QY 4201 GTTCGTATTCAGTAGTCAATTTGCCAGCATCAATTAATGGGATTAATCCAGAGCAACAG 4260  
 Db 3995 GTTCGTATTCAGTAGTCAATTTGCCAGCATCAATTAATGGGATTAATCCAGAGCAACAG 4054  
 QY 4261 TGGAGTCAACATCTACCAACTTTGGGTCTCAGAAAAGCATTAACAGTTCTACTACGC 4320  
 Db 4055 TGGAGTCAACATCTACCAACTTTGGGTCTCAGAAAAGCATTAACAGTTCTACTACGC 4114  
 QY 4321 CATTAGTCAAACTTTTCAATCGCCAGTGGAGAGAAAAGGACAGCATACTAGCAT 4380  
 Db 4115 CATTAGTCAAACTTTTCAATCGCCAGTGGAGAGAAAAGGACAGCATACTAGCAT 4174  
 QY 4381 TAGCGGCAAGATGCAACTTTATCAACAGGTCCTATAGATAACCTAGGCTCGGA 4440  
 Db 4175 TAGCGGCAAGATGCAACTTTATCAACAGGTCCTATAGATAACCTAGGCTCGGA 4234  
 QY 4441 TCATCCTTTGACACACTCTTTCTGCAATCTAGTCCAAATCACTTCATTGATACCAT 4500  
 Db 4235 TCATCCTTTGACACACTCTTTCTGCAATCTAGTCCAAATCACTTCATTGATACCAT 4294  
 QY 4501 TATTGTACAACTTTGACCAAGTTGTCAGTCTCTCAATTTGGTCTCTCTTAAGGATG 4560  
 Db 4295 TATTGTACAACTTTGACCAAGTTGTCAGTCTCTCAATTTGGTCTCTCTTAAGGATG 4354  
 QY 4561 ACTCAACTTGCACATTAATTTGAGCTCAGTCGATTGAGTGAATCTTGAATCAGTTGTGCA 4620  
 Db 4355 ACTCAACTTGCACATTAATTTGAGCTCAGTCGATTGAGTGAATCTTGAATCAGTTGTGCA 4414  
 QY 4621 GCTGTGACAGCATAGGAAAACACGGCTTTTCTPACAAAACCTCAAGAAATTAACAAT 4680  
 Db 4415 GCTGTGACAGCATAGGAAAACACGGCTTTTCTPACAAAACCTCAAGAAATTAACAAT 4474  
 QY 4681 CTGCAACACTTTCGCTATGAGTATGAGGAGGAAATGTCATCTTGAAGTTCGAGAGTGA 4740  
 Db 4475 CTGCAACACTTTCGCTATGAGTATGAGGAGGAAATGTCATCTTGAAGTTCGAGAGTGA 4534  
 QY 4741 TGTAGTCTTGAATAATCTGAAGCGGTATTTTATATCAGTGAAGTCAATCAGGAGA 4800  
 Db 4535 TGTAGTCTTGAATAATCTGAAGCGGTATTTTATATCAGTGAAGTCAATCAGGAGA 4594  
 QY 4801 TCTCTACCGCGAGCATCTGTCGCGA----- 4828  
 Db 4595 TCTCTACCGCGAGCATCTGTCGCGA----- 4654  
 QY 4829 ----- 4828  
 Db 4655 CTCGTGAAGAGGTGTTGCTGACTCATACAGGCTGAATCGCCCCATCATCCAGCCAGA 4714  
 QY 4829 ----- 4828



QY 5900 CAGACGGTCAAGCTTGTCTTAAGCGGATGCCGGGAGCAGACAAGCCCGCTCAGGGCGCG 5959  
DB 6935 CAGACGGTCAAGCTTGTCTTAAGCGGATGCCGGGAGCAGACAAGCCCGCTCAGGGCGCG 6994  
QY 5960 TCAGCGGGTGTGGCGGGTGTGGCGGGCGCAGCCATGACCCAGTCAAGTACGATAGCGGA 6019  
DB 6995 TCAGCGGGTGTGGCGGGTGTGGCGGGCGCAGCCATGACCCAGTCAAGTACGATAGCGGA 7054  
QY 6020 GTGTATACCTGCTTAACTATCCGCGCATCAGACGAGATTGTACTGAGAGTGCACCATATGC 6079  
DB 7055 GTGTATACCTGCTTAACTATCCGCGCATCAGACGAGATTGTACTGAGAGTGCACCATATGC 7114  
QY 6080 GGTGTGAATACCGCACAGATGCGTAAAGAGAAATACCGCATCAGCGGCTCTTCGCGTT 6139  
DB 7115 GGTGTGAATACCGCACAGATGCGTAAAGAGAAATACCGCATCAGCGGCTCTTCGCGTT 7174  
QY 6140 CCTCGCTCACTGACTCGCTGCGCTCGCTGCTGCTGCGCGAGCGGTATCAGCTCACT 6199  
DB 7175 CCTCGCTCACTGACTCGCTGCGCTCGCTGCTGCTGCGCGAGCGGTATCAGCTCACT 7234  
QY 6200 CAAAGGCGGTAAATACGGTATCCACAGAAATCAGGGGATAACGCAAGGAAAGAAATGTGAG 6259  
DB 7235 CAAAGGCGGTAAATACGGTATCCACAGAAATCAGGGGATAACGCAAGGAAAGAAATGTGAG 7294  
QY 6260 CAAAGGCGCAAAAGCCAGGAACCGTAAAGAGCGCGGTGCTGCGGTTTTCCATA 6319  
DB 7295 CAAAGGCGCAAAAGCCAGGAACCGTAAAGAGCGCGGTGCTGCGGTTTTCCATA 7354  
QY 6320 GGCTCCGCGCCCTGACGAGCATCACAAAATCAGCGCTCAAGTCAGAGGTGCGGAAACC 6379  
DB 7355 GGCTCCGCGCCCTGACGAGCATCACAAAATCAGCGCTCAAGTCAGAGGTGCGGAAACC 7414  
QY 6380 CGACAGGACTAAAGATACAGCGGTTTCCCGTGGAAAGCTCCCTGTCGCGCTCTCCTG 6439  
DB 7415 CGACAGGACTAAAGATACAGCGGTTTCCCGTGGAAAGCTCCCTGTCGCGCTCTCCTG 7474  
QY 6440 TTCGAGCCCTCCGCTTACCGGATACCTGCGGCTTTCTCCCTTCGGGAAGCGTGGCGC 6499  
DB 7475 TTCGAGCCCTCCGCTTACCGGATACCTGCGGCTTTCTCCCTTCGGGAAGCGTGGCGC 7534  
QY 6500 TTTCTCATPAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAGCTG 6559  
DB 7535 TTTCTCATPAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAGCTG 7594  
QY 6560 GCTGTGTGACGACCCCGCTTACCGCCGAGCCGCTGCGCTTATCCGGTAACTATCGTC 6619  
DB 7595 GCTGTGTGACGACCCCGCTTACCGCCGAGCCGCTGCGCTTATCCGGTAACTATCGTC 7654  
QY 6620 TTGAGTCCAAACCCGGTAAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA 6679  
DB 7655 TTGAGTCCAAACCCGGTAAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA 7714  
QY 6680 TTAGCAGACGAGGTATGAGCGGTGCTACAGAGTTCCTGAGTGTGCGCTTAACTACG 6739  
DB 7715 TTAGCAGACGAGGTATGAGCGGTGCTACAGAGTTCCTGAGTGTGCGCTTAACTACG 7774  
QY 6740 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTCACCTTCGAA 6799  
DB 7775 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTCACCTTCGAA 7834  
QY 6800 AAAGAGTTGGTGTCTTGTATCCGGCAAAACAAACCAACCGCTGTAGCGGTGTTTTT 6859  
DB 7835 AAAGAGTTGGTGTCTTGTATCCGGCAAAACAAACCAACCGCTGTAGCGGTGTTTTT 7894  
QY 6860 TTTCAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGTATCTTT 6919  
DB 7895 TTTCAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGTATCTTT 7954  
QY 6920 CTACGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAAGGATTTTGGTCAATGAGAT 6979  
DB 7955 CTACGGGTCTCAGCTCAGTGGAAACGAAACTCAGCTTAAAGGATTTTGGTCAATGAGAT 8014  
QY 6980 TATCAAAAAGGATCTTACCTAGATCCCTTTTAAATTAAGATGAAGTTTTAAATCAATCT 7039

DB 8015 TATCAAAAAGGATCTTACCTAGATCCCTTTTAAATTAAGATGAAGTTTTAAATCAATCT 8074  
QY 7040 AAAGTATATATGAGTAAACTTGGTCTGACAGTACCAGTCTTAATCAGTCAGGACACTA 7099  
DB 8075 AAAGTATATATGAGTAAACTTGGTCTGACAGTACCAGTCTTAATCAGTCAGGACACTA 8134  
QY 7100 TCTCAGCGATCTGTCTATTTCTGTTTATCCATAGTTCGCTGACTCCCGTCTGTAGATAA 7159  
DB 8135 TCTCAGCGATCTGTCTATTTCTGTTTATCCATAGTTCGCTGACTCCCGTCTGTAGATAA 8194  
QY 7160 CTACGATACGGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCGAGCCAC 7219  
DB 8195 CTACGATACGGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCGAGCCAC 8254  
QY 7220 GCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGCGAGAA 7279  
DB 8255 GCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGCGAGAA 8314  
QY 7280 GTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAATGTTGCGGGAAGTAGAG 7339  
DB 8315 GTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAATGTTGCGGGAAGTAGAG 8374  
QY 7340 TAAAGTAGTTGCGCAGTAAATAGTTTGCACAACTGTTGCCATTTGCTGACGGCATCTG 7399  
DB 8375 TAAAGTAGTTGCGCAGTAAATAGTTTGCACAACTGTTGCCATTTGCTGACGGCATCTG 8434  
QY 7400 TGTCAAGTGTGCTGTTGGTATGCTTCAATCAGTTCGGTTCCTCAAGCATCAAGGCGAG 7459  
DB 8435 TGTCAAGTGTGCTGTTGGTATGCTTCAATCAGTTCGGTTCCTCAAGCATCAAGGCGAG 8494  
QY 7460 TTACATGATCCCCATGTTGTGCAAAAAGCGTTAGCTCTCGTCTCCGATCGTTG 7519  
DB 8495 TTACATGATCCCCATGTTGTGCAAAAAGCGTTAGCTCTCGTCTCCGATCGTTG 8554  
QY 7520 TCAGAAGTAAGTTGCGCAGTGTATCACTCATGTTTATGCGCAGCACTGATATTTCTC 7579  
DB 8555 TCAGAAGTAAGTTGCGCAGTGTATCACTCATGTTTATGCGCAGCACTGATATTTCTC 8614  
QY 7580 TTACTGTATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCA 7639  
DB 8615 TTACTGTATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCA 8674  
QY 7640 TCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTCTTTCGCGCGCTCAACCGGATAATA 7699  
DB 8675 TCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTCTTTCGCGCGCTCAACCGGATAATA 8734  
QY 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAAGAGTCTTCGGGGCGAA 7759  
DB 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAAGAGTCTTCGGGGCGAA 8794  
QY 7760 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 7819  
DB 8795 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 8854  
QY 7820 ACTGATCTTCAGATCTTTTACTTTTCAACGCTTTCGGGTGAGCAAAAACAGAGGCG 7879  
DB 8855 ACTGATCTTCAGATCTTTTACTTTTCAACGCTTTCGGGTGAGCAAAAACAGAGGCG 8914  
QY 7880 AAAATGCGCAAAAAGGAAATAGGGCGACACGGAATTTGAAATCTCATCTCTTCC 7939  
DB 8915 AAAATGCGCAAAAAGGAAATAGGGCGACACGGAATTTGAAATCTCATCTCTTCC 8974  
QY 7940 TTTTTCATATTTATGAAGCATTTATCAGGTTTATGTCTCATGAGCGGATACATATTTG 7999  
DB 8975 TTTTTCATATTTATGAAGCATTTATCAGGTTTATGTCTCATGAGCGGATACATATTTG 9034  
QY 8000 AATGATTTAGAAAATAAACAAATAGGGTTCCGCGACATTTCCCGAAAAGTGCCAC 8059  
DB 9035 AATGATTTAGAAAATAAACAAATAGGGTTCCGCGACATTTCCCGAAAAGTGCCAC 9094  
QY 8060 CTGAGCTCTAAGAAACCACTTATTATCATGACATTAACCTATATAAATAGGCGTATCAGA 8119

Db 9095 CTGACGCTTAAGAAACCATTTATTATCATGACATTAACTATATAAAATAGCGGTATCACGA 9154

Qy 8120 GGCCTCTTGGCTTTCAAGAAATTAATTTCTCATGTTTGAAGCTTATCATCGATAGCTGAC 8179

Db 9155 GGCCTCTTGGCTTTCAAGAAATTAATTTCTCATGTTTGAAGCTTATCATCGATAGCTGAC 9214

Qy 8180 TCATGTTGGTATTGTAATAGACGAGATCGGAAACACATGAAAAATAACAGTTATTAT 8239

Db 9215 TCATGTTGGTATTGTAATAGACGAGATCGGAAACACATGAAAAATAACAGTTATTAT 9274

Qy 8240 CG 8241

Db 9275 CG 9276

RESULT 9

LOCUS AX146618 9776 bp DNA linear PAT 31-MAY-2001

DEFINITION Sequence 80 from Patent WO0134654.

ACCESSION AX146618

VERSION AX146618.1 GI:14285011

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1. Strauch, K.

AUTHORS

TITLE Hedgehog fusion proteins and uses

JOURNAL Patent: WO 0134654-A 80 17-MAY-2001;

BIOGEN, INC. (US)

FEATURES

source 1.9776

location/Qualifiers

1.9776

organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 77.5%; Score 6383.8; DB 6; Length 9776;

Best Local Similarity 82.9%; Pred No. 0;

Matches 8101; Conservative 0; Mismatches 137; Indels 1535; Gaps 10;

Qy 2 GATCTAACATCCAAAGAGCAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 61

Db 1 GATCTAACATCCAAAGAGCAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 60

Qy 62 TCATCTCAGACATAGTGCACCAAGCAAGAGGGGATACACTAGCAGACAGCGTT 121

Db 61 TCATCTCAGACATAGTGCACCAAGCAAGAGGGGATACACTAGCAGACAGCGTT 120

Qy 122 GCAAGCGAGGACCTCCACTCTCTCTCTCAACACCCACTTTTGGCCATCGAAAAACCA 181

Db 121 GCAAGCGAGGACCTCCACTCTCTCTCTCAACACCCACTTTTGGCCATCGAAAAACCA 180

Qy 182 GCCAGTTATGGGCTTGATTTGGAGTGGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 241

Db 181 GCCAGTTATGGGCTTGATTTGGAGTGGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 240

Qy 242 CACCATGCTTTATAGCTGCTCTATCTCGGCCCTTGGCGAGGTTTATGTTTCTTTAT 301

Db 241 CACCATGCTTTATAGCTGCTCTATCTCGGCCCTTGGCGAGGTTTATGTTTCTTTAT 300

Qy 302 TTCGAATCAACAAGCTCCGATTAACCCGAAACATCACTCCAGATGAGGCTTTCTGA 361

Db 301 TTCGAATCAACAAGCTCCGATTAACCCGAAACATCACTCCAGATGAGGCTTTCTGA 360

Qy 362 GTGTGGGTTCAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAACGCTG 421

Db 361 GTGTGGGTTCAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAACGCTG 420

Qy 422 TCTTGAACCTTAATGACAAAGCGTGATCTCATCCAAAGATGAACCTTAAGTTTGGTTCT 481

Db 421 TCTTGAACCTTAATGACAAAGCGTGATCTCATCCAAAGATGAACCTTAAGTTTGGTTCT 480

Qy 482 TGAATGCTAAGCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCCATACCGTTTCTC 541

Db 481 TGAATGCTAAGCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCCATACCGTTTCTC 540

Qy 542 TTGTTTGGTATTGATGACGAATGCTCAAAAATAATCTCATTAATCTTAGCGAGTCTC 601

Db 541 TTGTTTGGTATTGATGACGAATGCTCAAAAATAATCTCATTAATCTTAGCGAGTCTC 600

Qy 602 TCTATCGCTTCTGAACCCCGGTGCACTGTCGCGAAACGCAATGGGAAACACCCGCTT 661

Db 601 TCTATCGCTTCTGAACCCCGGTGCACTGTCGCGAAACGCAATGGGAAACACCCGCTT 660

Qy 662 TTTGGATGATTATGATGCTCTCCATGTTGATGCTTCCAAAGTCTGTTGGGAATACTG 721

Db 661 TTTGGATGATTATGATGCTCTCCATGTTGATGCTTCCAAAGTCTGTTGGGAATACTG 720

Qy 722 CTGATAGCTTAACGTTTATGATCAAAATTAATTAATCTTAACCCCTACTTGACGAATA 781

Db 721 CTGATAGCTTAACGTTTATGATCAAAATTAATTAATCTTAACCCCTACTTGACGAATA 780

Qy 782 TATAACAGAGGAGTGCCTGCTTAAACCTTTTATCATCATTAATTAAGCTTA 841

Db 781 TATAACAGAGGAGTGCCTGCTTAAACCTTTTATCATCATTAATTAAGCTTA 840

Qy 842 CTTTCAATAATTCGAGTCTGTTTCAATTTGAAGCTTTTGAATTTAAACGCTTTTAAACGAC 901

Db 841 CTTTCAATAATTCGAGTCTGTTTCAATTTGAAGCTTTTGAATTTAAACGCTTTTAAACGAC 900

Qy 902 AACTGAGAGATCAAAAACAACTTAATTTTCGAGGATCCAAAGATGAGATTTCCCTT 961

Db 901 AACTGAGAGATCAAAAACAACTTAATTTTCGAGGATCCAAAGATGAGATTTCCCTT 960

Qy 962 CAATTTTACTGAGTCTTATTCGAGCATCTCCGATTTAGCTGCTCCAGTCAACACTA 1021

Db 961 CAATTTTACTGAGTCTTATTCGAGCATCTCCGATTTAGCTGCTCCAGTCAACACTA 1020

Qy 1022 CAACAGAGATGAACCGCACAAATTCGGCTGGAAGCTGTCATCGCTTACTCAGATTAG 1081

Db 1021 CAACAGAGATGAACCGCACAAATTCGGCTGGAAGCTGTCATCGCTTACTCAGATTAG 1080

Qy 1082 AAGGGATTTGATGTTGCTGTTTTCGCAATTTTCCAAAGCAGCAAAATAACGGGTATTGT 1141

Db 1081 AAGGGATTTGATGTTGCTGTTTTCGCAATTTTCCAAAGCAGCAAAATAACGGGTATTGT 1140

Qy 1142 TTATAAATACTACTATTGCCAGCATTTGCTGCTAAAGAAAGAGGGTATCTCTGAGAAAA 1201

Db 1141 TTATAAATACTACTATTGCCAGCATTTGCTGCTAAAGAAAGAGGGTATCTCTGAGAAAA 1200

Qy 1202 GAGAGCTGAGCCCGAGG----- 1219

Db 1201 GATCGGACCGGCGAGGGGTTCCGGAAGAGGAGGCCACCCAAAAGCTGACCCCTTTAG 1260

Qy 1220 -----CCAGGCCCGAGGCCCGAG 1242

Db 1261 CCTACAGCAGTTTATCCCAATGTGGCCGAGAGAGCCCTAGCGCCGAGGAGGTATG 1320

Qy 1243 GAAGAAACAT----- 1252

Db 1321 AAGGGAAGATCTCCAGAACTCCGAGCGATTTAAGGAATCAACCCCAATTAACAACCCG 1380

Qy 1253 -----GTATCATGGCCCCCGGAGAGGATAAATTTGTGGCTTCCCGGTGTC 1299

Db 1381 ACATCATATTTAAGGATGAAGAAACACCGAGCGGACAGGCTGATGACTCAGAGGTGTA 1440

Qy 1300 AC-----CGCCAGCAGTGCACGAGAGAGGTTGCTGTTT 1334

Db 1441 AGGACAAAGTTGAACCGCTTTGGCCATCTCGGTGATGAACAGTGGCCAGGAGTGAACCTGC 1500

Qy 1335 TGATGACAGTGTCCGGGATTCGGTG----- 1362

Db 1501 GGGTGACCGAGGCTGGGACGAAGATGGCCACCACTCAGAGGAGTCTCTGCACCTACAGG 1560

QY 1363 ----- 1362  
Db 1561 GCCGGCAGTGGACATCACACAGTCTGACCGGACCGCAGCAAGTACGGCATGCTGGCCC 1620  
QY 1363 -----TGCTTCCACCCCATGGCCATCGAGAACACACTCAAGAAGAAGAT 1405  
Db 1621 GCCTGGCGGTGGAGCCGGCTTCAGCTGGGTACTACAGGTCCAGGTCCAGGCACATATCCACT 1680  
QY 1406 GTCCCTTCTACTA-----GTGGGTAGANTCCCTAGG-----GCGGCGCGGA 1449  
Db 1681 GCTCGGTGAAGCAGAGAACTCGGTGGCGGCCAAATCGGAGGCTGATTCGCGGCCCGGA 1740  
QY 1450 ATTAATTCGCCCTTAGACATGACTCTTCTCAGTTCAAGTTTGGGCACCTTACGAGAAGACCG 1509  
Db 1741 ATTAATTCGCCCTTAGACATGACTCTTCTCAGTTCAAGTTTGGGCACCTTACGAGAAGACCG 1800  
QY 1510 GTCTTGCTAGATTCTAATCAAGAGATGTGAGATGCAATTTGCTCGAGAGATGCAGGCT 1569  
Db 1801 GTCTTGCTAGATTCTAATCAAGAGATGTGAGATGCAATTTGCTCGAGAGATGCAGGCT 1860  
QY 1570 TCATTTTGTACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTG 1629  
Db 1861 TCATTTTGTACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTG 1920  
QY 1630 TTTCTTCTCGTACGAGCTTGCTCTGATCAGCTTATCTCGAGCTGATGAATATCTTGTG 1689  
Db 1921 TTTCTTCTCGTACGAGCTTGCTCTGATCAGCTTATCTCGAGCTGATGAATATCTTGTG 1980  
QY 1690 GTAGGGTTTGGGAAATCATTTCAGTTTGTATGTTTTTCTTGTGATTTCCACTCCTCTT 1749  
Db 1981 GTAGGGTTTGGGAAATCATTTCAGTTTGTATGTTTTTCTTGTGATTTCCACTCCTCTT 2040  
QY 1750 CAGAGTACAGAAGATTAAGTGAGAAGTTGTTTGTGCAAGCTTATCGATAAGCTTTAATG 1809  
Db 2041 CAGAGTACAGAAGATTAAGTGAGAAGTTGTTTGTGCAAGCTTATCGATAAGCTTTAATG 2100  
QY 1810 CGGTAGTTTATCAGAGTTAAATCTTACCGAGTCAGGACCGGTGATGAATCTAACAA 1869  
Db 2101 CGGTAGTTTATCAGAGTTAAATCTTACCGAGTCAGGACCGGTGATGAATCTAACAA 2160  
QY 1870 TCGCTCATCTGCTATCTCGGCACCGCTCACCCCTGGATGCTGTAGGCATAGGCTTGTAT 1929  
Db 2161 TCGCTCATCTGCTATCTCGGCACCGCTCACCCCTGGATGCTGTAGGCATAGGCTTGTAT 2220  
QY 1930 GCCGTACTGCCGGCTCTTTGGGGATATGTCCTATTCGACAGCATGCCAGTCACTA 1989  
Db 2221 GCCGTACTGCCGGCTCTTTGGGGATATGTCCTATTCGACAGCATGCCAGTCACTA 2280  
QY 1990 TGGCTGTCTAGCGCTATATGTTGTATGATCAATTTCTATGCCACCCGCTTCTCGGAGC 2049  
Db 2281 TGGGTGTCTAGCGCTATATGTTGTATGATCAATTTCTATGCCACCCGCTTCTCGGAGC 2340  
QY 2050 ACTGTCCGACCGCTTTGGCGCCGCCAGTCTGCTCGTTTGTCTATTCGGAGCCACTAT 2109  
Db 2341 ACTGTCCGACCGCTTTGGCGCCGCCAGTCTGCTCGTTTGTCTATTCGGAGCCACTAT 2400  
QY 2110 CGACTACGATATGCGGACCAACACCCGCTCTGTGGATCTATCGAATCTAAATGTAAAGT 2169  
Db 2401 CGACTACGATATGCGGACCAACACCCGCTCTGTGGATCTATCGAATCTAAATGTAAAGT 2460  
QY 2170 TAAATCTCTAAATTAATAATAGTCCCAAGTTTCTCCATACGAACCTTAAACAGCATTC 2229  
Db 2461 TAAATCTCTAAATTAATAATAGTCCCAAGTTTCTCCATACGAACCTTAAACAGCATTC 2520  
QY 2230 GGTGAGATCTAGACCTTCAACAGCAGCAGATCATCTGCTTGGCAATATGTTTCA 2289  
Db 2521 GGTGAGATCTAGACCTTCAACAGCAGCAGATCATCTGCTTGGCAATATGTTTCA 2580  
QY 2290 GTCCCTCAGGATTTACGTTCTTGTGAAGTGTAACTTTCTGGAAGTTTGCAGTGTAACTC 2349  
Db 2581 GTCCCTCAGGATTTACGTTCTTGTGAAGTGTAACTTTCTGGAAGTTTGCAGTGTAACTC 2640  
QY 2350 CGCTGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGGTACCGAGGAGTAA 2409

Db 2641 CGCTGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGGTGTACCGGAGGTAAT 2700  
QY 2410 CTCCTCAACTCTCTCTGGAGTAGGACCAACAAACACAGATCCAGCGTGTGTACTTGAT 2469  
Db 2701 CTCCTCAACTCTCTCTGGAGTAGGACCAACAAACACAGATCCAGCGTGTGTACTTGAT 2760  
QY 2470 CAACATAGAAGAAGCATTCTCGATTTGCAGATCAAGTGTTCAGGAGCGTACTGATGG 2529  
Db 2761 CAACATAGAAGAAGCATTCTCGATTTGCAGATCAAGTGTTCAGGAGCGTACTGATGG 2820  
QY 2530 ACATTTCCAAAGCCTGTCTGATGTTGCAACCGATAGGGTTGTAGAGTGTGCAATACACT 2589  
Db 2821 ACATTTCCAAAGCCTGTCTGATGTTGCAACCGATAGGGTTGTAGAGTGTGCAATACACT 2880  
QY 2590 TCGGTACAAATTTCAACCCCTTGGCAACTGCAACGCTTGTGTGAACAGCATCTTCAATTC 2649  
Db 2881 TCGGTACAAATTTCAACCCCTTGGCAACTGCAACGCTTGTGTGAACAGCATCTTCAATTC 2940  
QY 2650 TGGCAAGCTCTCTTGTCTGTCTATATCGACAGCCAAACAGAAATCACCTGGGAATCAATACCAT 2709  
Db 2941 TGGCAAGCTCTCTTGTCTGTCTATATCGACAGCCAAACAGAAATCACCTGGGAATCAATACCAT 3000  
QY 2710 GTTCAGCTTGACACAGAGTCTGAGGCAACGAATCTGGATCAGCGTATTTATCAGCAA 2769  
Db 3001 GTTCAGCTTGAG-CAGAAAGTCTGAGGCAACGAATCTGGATCAGCGTATTTATCAGCAA 3059  
QY 2770 TAACTAGAACTTCAGAAAGCCAGCAGCATGTCAATACTACACAGGCGCTGATGTCTAT 2829  
Db 3060 TAACTAGAACTTCAGAAAGCCAGCAGCATGTCAATACTACACAGGCGTATGTCTAT 3119  
QY 2830 TTTGAAACATCATCTTGGCAGCATGCAACTGGTTTCTTGGACCAAAATTTTGTCTAC 2889  
Db 3120 TTTGAAACATCATCTTGGCAGCATGCAACTGGTTTCTTGGACCAAAATTTTGTCTAC 3179  
QY 2890 ACTTAGGAAACAGTTTCTGTTCGTAAAGCATAGCAGCTACTGCTTGGCGGCTCTCTGCTA 2949  
Db 3180 ACTTAGGAAACAGTTTCTGTTCGTAAAGCATAGCAGCTACTGCTTGGCGGCTCTCTGCTA 3239  
QY 2950 GCACGATACACTTAGCACCAACCTTGTGGCAACGATAGATCACTTCTGGGTAGGGTAC 3009  
Db 3240 GCACGATACACTTAGCACCAACCTTGTGGCAACGATAGATCACTTCTGGGTAGGGTAC 3299  
QY 3010 CATCTCTTCTTAGGTGAGATGCAAAACAAATTTCTTTCACACAGCACTTTTGCAGGAA 3069  
Db 3300 CATCTCTTCTTAGGTGAGATGCAAAACAAATTTCTTTCACACAGCACTTTTGCAGGAA 3359  
QY 3070 CACCAGCATCAGGGAAGTGGAAAGCGAATTCGGGTTCCACCGAATATAGAGGCCAA 3129  
Db 3360 CACCAGCATCAGGGAAGTGGAAAGCGAATTCGGGTTCCACCGAATATAGAGGCCAA 3419  
QY 3130 CTTTCTCAATAGGTTTCCAAACAGAGCAGACTACACAGGCGAAGTCTCACTTGCA 3189  
Db 3420 CTTTCTCAATAGGTTTCCAAACAGAGCAGACTACACAGGCGAAGTCTCACTTGCA 3479  
QY 3190 AGCTCTCCGTTAGTTAGCTTCATGGATTTCTGAGCTTATCTATAGAGAGTCAATGG 3249  
Db 3480 AGCTCTCCGTTAGTTAGCTTCATGGATTTCTGAGCTTATCTATAGAGAGTCAATGG 3539  
QY 3250 CTTCTTCAAGTATCTGGCAATTCGATAGTTCCTCTGGGAAAGAGCTTCTAACACAG 3309  
Db 3540 CTTCTTCAAGTATCTGGCAATTCGATAGTTCCTCTGGGAAAGAGCTTCTAACACAG 3599  
QY 3310 GTGTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCGCTTTGTCCCAT 3369  
Db 3600 GTGTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCGCTTTGTCCCAT 3659  
QY 3370 TTTGACGAACATGTGCAAAATTTGTTTGAATAATCCATATCTGTTCGTTTTTGGGA 3429  
Db 3660 TTTGACGAACATGTGCAAAATTTGTTTGAATAATCCATATCTGTTCGTTTTTGGGA 3719  
QY 3430 TAGGACGACGAGGCGCTTCAATTTCTTGTAGGAGGCGCTTAGAAACGTCATTTTTCG 3489



Db	3720	TAGGACGACGAAGGGCAATCTTCAAATTTCTTGAGAGAGCCCTTAGAAACGTCOAATTTTGC	3779
Qy	3490	ACAATTAACAATACGACCTTCAGAAAGGACACTCTTTTAGTTTGGATTCTCTCTTTAGGTGTT	3549
Db	3780	ACAATTAACAATACGACCTTCAGAAAGGACACTCTTTAGTTTGGATTCTCTTTAGGTGTT	3839
Qy	3550	CCTTGGTGTATCCTGGCTTGGCATTCTCCTTCTAGTGACCTTTAGGACCTTCATAT	3609
Db	3840	CCTTGGTGTATCCTGGCTTGGCATTCTCCTTCTAGTGACCTTTAGGACCTTCATAT	3899
Qy	3610	CCAGGTTTCTCTCCACTCGTCCAAAGCTCACACCGTACTTGGCACATCTAACTAATGCAA	3669
Db	3900	CCAGGTTTCTCTCCACTCGTCCAAAGCTCACACCGTACTTGGCACATCTAACTAATGCAA	3959
Qy	3670	AATAAATAAGTCAGACACATTCACAGCTATATCTTCTCTTGGATTTAGCTTCTGCAAGTT	3729
Db	3960	AATAAATAAGTCAGACACATTCACAGCTATATCTTCTCTTGGATTTAGCTTCTGCAAGTT	4019
Qy	3730	CATCAGCTTCCCTCCCTAAATTTAGCGTTC-AACAAAACTTCGTCGTCAAATAACCGTTTG	3788
Db	4020	CATCAGCTTCCCTCCCTAAATTTAGCGTTCAAACAAAATTCGTCGTCAAATAACCGTTTG	4079
Qy	3789	GTATAAGAACCTTCTGGAGCATTTGCTTTACGATCCCAAGTCGCTTCCATGGCTCTA	3848
Db	4080	GTATAAGAACCTTCTGGAGCATTTGCTTTACGATCCCAAGT-CTTTCCTAGTGGCTCTA	4138
Qy	3849	AGACCTTTGATTGGCCAAACAGGAGTGCCTTCCAAGTGCACAGAAAAACAACCTGTT	3908
Db	4139	AGACCTTTGATTGGCCAAACAGGAGTGCCTTCCAAGTGCACAGAAAAACAACCTGTT	4198
Qy	3909	TGTTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACCCAGCAACTTT	3968
Db	4199	TGTTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACCCAGCAACTTT	4258
Qy	3969	TGAGTTGCTCCAGATGTAGCACCTTTATACCAAAAACCGTGACGACGAGTTGGTAGCT	4028
Db	4259	TGAGTTGCTCCAGATGTAGCACCTTTATACCAAAAACCGTGACGACGAGTTGGTAGCT	4318
Qy	4029	CCAGTTTGTCCTTTAGCCTTCGGAAATAGACTTTTTTGACGAGTACACACGAGCCCAAC	4088
Db	4319	CCAGTTTGTCCTTTAGCCTTCGGAAATAGACTTTTTTGACGAGTACACACGAGCCCAAC	4378
Qy	4089	GAGTAATTAGAGAGTCAAGCCACCAAAAGTAGTGAAATAGACCATCGGGGGCGGTACGTAGTC	4148
Db	4379	GAGTAATTAGAGAGTCAAGCCACCAAAAGTAGTGAAATAGACCATCGGGGGCGGTACGTAGTC	4438
Qy	4149	AAAGACGCCAACAAAATTTCACTGACAGGGAACTTTTTGACATCTTCAGAAAAGTTCGTAT	4208
Db	4439	AAAGACGCCAACAAAATTTCACTGACAGGGAACTTTTTGACATCTTCAGAAAAGTTCGTAT	4498
Qy	4209	TCAGTAGTCAATTCGAGCATCAATTAATGGGANTTATACAGAGCAACAGTGGAGTGC	4268
Db	4499	TCAGTAGTCAATTCGAGCATCAATTAATGGGANTTATACAGAGCAACAGTGGAGTGC	4558
Qy	4269	ACATCTACCAACTTTGGCGTCTCAGAAAAAGCATAAACAGTTCTACTACCGCCATTAGTG	4328
Db	4559	ACATCTACCAACTTTGGCGTCTCAGAAAAAGCATAAACAGTTCTACTACCGCCATTAGTG	4618
Qy	4329	AAACTTTTCAAATCGCCAGTGGAGAGAAAAAGGCACAGCATCTAGCATTTAGCGGGC	4388
Db	4619	AAACTTTTCAAATCGCCAGTGGAGAGAAAAAGGCACAGCATCTAGCATTTAGCGGGC	4678
Qy	4399	AAGGATGCAACTTTATCAACAGGGTCTATAGATACCCCTAGCGCTGGGATCATCCTT	4448
Db	4679	AAGGATGCAACTTTATCAACAGGGTCTATAGATACCCCTAGCGCTGGGATCATCCTT	4738
Qy	4449	TGGACCACTCTTTCTGCCAAATCTAGGTCCAAAAATCACTTCATTGATACCATTTGTAC	4508
Db	4739	TGGACCACTCTTTCTGCCAAATCTAGGTCCAAAAATCACTTCATTGATACCATTTGTAC	4798
Qy	4509	AACHTGACAAAGTTGTGATCAGCTCCTCAAAATTTGGTCTCTGTGAAACGGATGACTCACT	4568
Db	4799	AACHTGACAAAGTTGTGATCAGCTCCTCAAAATTTGGTCTCTGTGAAACGGATGACTCACT	4858

QY	4569	TGCACATTAACTGAAAGCTCAGTCGATTGAGTGAACCTTGATCAGGTTGTGCAGCTGGTCA	4628
DB	4859	TGCACATTAACTGAAAGCTCAGTCGATTGAGTGAACCTTGATCAGGTTGTGCAGCTGGTCA	4918
QY	4629	GCAGCATAGGAAACACGGCTTTTCCCTACCAAACTCAAGGAATTATCAAACTCTGCAACA	4688
DB	4919	GCAGCATAGGAAACACGGCTTTTCCCTACCAAACTCAAGGAATTATCAAACTCTGCAACA	4978
QY	4689	CTTGCCTATGTCAGGTAGCAAGGGAATGTCATCTTGAAGTCGGAAGTGTAGTCT	4748
DB	4979	CTTGCCTATGTCAGGTAGCAAGGGAATGTCATCTTGAAGTCGGAAGTGTAGTCT	5038
QY	4749	TGAGAAATCTGAAGCCGTATTTTATTAATCACTAGTGAGTCAGTCATCAGCAGATCCCTCTAC	4808
DB	5039	TGAGAAATCTGAAGCCGTATTTTATTAATCACTAGTGAGTCAGTCATCAGCAGATCCCTCTAC	5098
QY	4809	GCCGGACGCATCGTGGCCGACCTGSCAGTCCGCATCACCGGGCCACAGGTGCGTTGCT	4868
DB	5099	GCCGGACGCATCGTGGCCGACCTGSCAGTCCGCATCACCGGGCCACAGGTGCGTTGCT	5158
QY	4869	GGGCGCTATATCGCGACATCACCGATGGGAAGATCGGGCTCGCCACTTCGGGCTCATG	4928
DB	5159	GGGCGCTATATCGCGACATCACCGATGGGAAGATCGGGCTCGCCACTTCGGGCTCATG	5218
QY	4929	AGGCTTGTTTCGGCGTGGGTATGTGTGGCAGGCCCGTGGCCGGGGACTGTGTGGCGCC	4988
DB	5219	AGGCTTGTTTCGGCGTGGGTATGTGTGGCAGGCCCGTGGCCGGGGACTGTGTGGCGCC	5278
QY	4989	ATCTCCTTG-----	4997
DB	5279	ATCTCCTTGACCTGCAGGGGGGGGGGAAAGCACCGTTGTCTCAAAATCTCTGAT	5338
QY	4998	-----	4997
DB	5339	GTTACATTGCACAGATAAAAATATATCATCATGAACATAAACTGTCTGCTTACATAA	5398
QY	4998	-----	4997
DB	5399	ACAGTAATCAAGGGTGTATTAGAGCATATTCAAACGGAAACGTCTTGCTCAAGGCCGC	5458
QY	4998	-----	4997
DB	5459	GATTAAATCCAACTGATGCTGATTTATGGGTATAAATGGGCTCGCGATAATGTCG	5518
QY	4998	-----	4997
DB	5519	GGCAATCAGGTGCGACAATCTATCGATTGTATGGAGCCCGATGGCCAGAGTTGTTTC	5578
QY	4998	-----	4997
DB	5579	TGAAACATGCGAAAGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAACT	5638
QY	4998	-----	4997
DB	5639	GGCTGACGAATTTATGCCTCTCCGACCATCAAGCATTTTATCCGTACTCTCTGATGATG	5698
QY	4998	-----	4997
DB	5699	CATGTTACTCACCACCTCGATCCCGGGAAACACAGATTCAGGTATTAGAGAATATC	5758
QY	4998	-----	4997
DB	5759	CTGATTACAGGTGAAAAATATTGTTGATGCGGTGGCAGTGTCTCTCGCGCGGTTCATTGCA	5818
QY	4998	-----	4997
DB	5819	TTCTGTGTTGTAATGTCTTTTAAACAGCGATCGCGTATTTCTGTCTCGCTCAGCGCAAT	5878
QY	4998	-----	4997
DB	5879	CACGAATGAATAACGGTTTGTGTCGAGTGATTTTGTATGACGAGCGTAATCGCTGGC	5938

QY 4998 ----- 4997  
Db 5939 CTGTTGAACAAGCTGGGAAAGAAATGCAATAGCTTTTGGCCATTTCTCACGGATTCAGTCG 5998  
QY 4998 ----- 4997  
Db 5999 TCACCTCATGGTGAATTTCTCACTTCGTATAACCTTATTTTGGACGAGGGGAAATTAATAGGTT 6058  
QY 4998 ----- 4997  
Db 6059 GTAATTGATGTTGGACGAGTCGGAATTCGACAGACCGATACAGGATCTTTGGCCATCCTATGGA 6118  
QY 4998 ----- 4997  
Db 6119 ACTGCCTCGGTGAGTTTCTCCTTCATTAACAGAACGGCTTTTCAAAAATATGGTATTG 6178  
QY 4998 ----- 4997  
Db 6179 ATAATCCTGATATGAATAAATTGCAGTTTCAATTTGATGCTCGATGAGTTTCTTAATACAG 6238  
QY 4998 ----- 4997  
Db 6239 AATTGGTTAATTGTTGTAACACTGGCAGACCATTAAGCTGACTTGACGGGACGGCGGCT 6298  
QY 4998 ----- 4997  
Db 6299 TTGTTGAATAAATCGAACTTTTGTGCTGAGTTGAAGGATCAGATCAGGCATCTTCCCGACAA 6358  
QY 4998 ----- 4997  
Db 6359 CGCAGACCTGTCGGTGGCAAGCAAAAGTTCAAAATCACCAATGCTGCTACAAACAA 6418  
QY 4998 ----- 4997  
Db 6419 AGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGCGGATTGAGGCGCTG 6478  
QY 4998 ----- 5008  
Db 6479 GTATGAGTCAGCAACACCTTTTCAAGAGGAGACCTCAGCGCCCCCCCCCTGCGAG 6538  
QY 5009 CTTTCCGCGCGGGTGTCAACGGCTCAACTACTACTGGCTGTCTTCTTAATGACAGGA 5068  
Db 6539 TCCACGCGCGGGTGTCAACGGCTCAACTACTACTGGCTGTCTTCTTAATGACAGGA 6598  
QY 5069 GTCCGATTAAGGGAGAGCGTCGAGTATCTATGATTTGGAAGTATGGGAATGGTATACCCGC 5128  
Db 6599 GTCCGATTAAGGGAGAGCGTCGAGTATCTATGATTTGGAAGTATGGGAATGGTATACCCGC 6658  
QY 5129 ATTCTTCAGTGTCTTGGGTCTCCTATCAGATTAATGCCAACTAAAGCAACCGGAGGAGG 5188  
Db 6659 ATTCTTCAGTGTCTTGGGTCTCCTATCAGATTAATGCCAACTAAAGCAACCGGAGGAGG 6718  
QY 5189 AGATTTCAATGTAATTTCTCTGACTTTTGTCTATCATGACTAGCTCGAACTGTGAGACTAT 5248  
Db 6719 AGATTTCAATGTAATTTCTCTGACTTTTGTCTATCATGACTAGCTCGAACTGTGAGACTAT 6778  
QY 5249 CTCGGTTATGACAGCAAGATGCTCTTCTGGAGACAGTAATGAAGTCCCAACCAATAAA 5308  
Db 6779 CTCGGTTATGACAGCAAGATGCTCTTCTGGAGACAGTAATGAAGTCCCAACCAATAAA 6838  
QY 5309 GAAATCTTGTATTCAGGAACAAACTTCTTGTTCGAACTTTTTCGGTGCCTTGAACATAT 5368  
Db 6839 GAAATCTTGTATTCAGGAACAAACTTCTTGTTCGAACTTTTTCGGTGCCTTGAACATAT 6898  
QY 5369 AAAATGATGATGATATGTCGGTGAAGATGGAGCGGCAATGCTTACTCTTGAGCC 5428  
Db 6899 AAAATGATGATGATATGTCGGTGAAGATGGAGCGGCAATGCTTACTCTTGAGCC 6958  
QY 5429 TTCAAGAGGTATGATGGGTTTGTAGATACTGATGCCAACTTTCAGTGACAACTGTGCTATT 5488  
Db 6959 TTCAAGAGGTATGATGGGTTTGTAGATACTGATGCCAACTTTCAGTGACAACTGTGCTATT 7018  
QY 5489 TCGTTCAAAACCAATTCGGAATCCAGAGAAATCAAAAGTTGTTGTCTACTATTGATCCAGC 5548

Db 7019 TCGTTCAAAACCAATTCGGAATCCAGAGAAATCAAAAGTTGTTGTCTACTATTGATCCAGC 7078  
QY 5549 CAGTGGCGTCTTGAACCTGCAATAGTGTGCTGCTGTTTGGAGTCACTTTTGTATGAAAT 5608  
Db 7079 CAGTGGCGTCTTGAACCTGCAATAGTGTGCTGCTGTTTGGAGTCACTTTTGTATGAAAT 7138  
QY 5609 AAATCTAGTCTTGTATCTAAATTAATCTTGAAGCAAGGCGATAAATACCAAAATCTAA 5668  
Db 7139 AAATCTAGTCTTGTATCTAAATTAATCTTGAAGCAAGGCGATAAATACCAAAATCTAA 7198  
QY 5669 AAATCTTTPAAAACCTTAAAGGCAAGTATGCTGCTGCTTAAACCCCAATACAGCT 5728  
Db 7199 AAATCTTTPAAAACCTTAAAGGCAAGTATGCTGCTGCTTAAACCCCAATACAGCT 7258  
QY 5729 CGTAGTCTGATCCTCATCAACTTGAAGGGGCACTATCTTGTGTTTAGAGAAATTTGCGGAGA 5788  
Db 7259 CGTAGTCTGATCCTCATCAACTTGAAGGGGCACTATCTTGTGTTTAGAGAAATTTGCGGAGA 7318  
QY 5789 TCGGATATCGAGAAAAGGTACGCTGATTTTAAACGCTGAAATTTTATCTCAAGATCTCTGC 5848  
Db 7319 TCGGATATCGAGAAAAGGTACGCTGATTTTAAACGCTGAAATTTTATCTCAAGATCTCTGC 7378  
QY 5849 CTCGGCGGTTTCGGTGATGACGGTGAAACCTCTGACATGCGACTCCCGGAGACGGTC 5908  
Db 7379 CTCGGCGGTTTCGGTGATGACGGTGAAACCTCTGACATGCGACTCCCGGAGACGGTC 7438  
QY 5909 ACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAAGCCGCTCAGGGCGCGCTCAGCGGT 5968  
Db 7439 ACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAAGCCGCTCAGGGCGCGCTCAGCGGT 7498  
QY 5969 GTTGGCGGTTTCGGGCGCAGCCATGACCCAGTCAAGTACGATAGCGAGTGTATCT 6028  
Db 7499 GTTGGCGGTTTCGGGCGCAGCCATGACCCAGTCAAGTACGATAGCGAGTGTATCT 7558  
QY 6029 GGCCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATCGGTGTGAAA 6088  
Db 7559 GGCCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATCGGTGTGAAA 7618  
QY 6089 TACCGCAGAGTGTCTAAGGAGAAATACCGCATCAGCGCTCTTTCGCTCTCTCGCTCA 6148  
Db 7619 TACCGCAGAGTGTCTAAGGAGAAATACCGCATCAGCGCTCTTTCGCTCTCTCGCTCA 7678  
QY 6149 CTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCTCAGTCAAAAGGCGG 6208  
Db 7679 CTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCTCAGTCAAAAGGCGG 7738  
QY 6209 TAATACGGTTATCCA CAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCC 6268  
Db 7739 TAATACGGTTATCCA CAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCC 7798  
QY 6269 AGCAAAAGCCAGGAAACCGTAAAGGCGCGTTCGCTGCGCTTTTCCATAGTCTCGGCC 6328  
Db 7799 AGCAAAAGCCAGGAAACCGTAAAGGCGCGTTCGCTGCGCTTTTCCATAGGCTCGGCC 7858  
QY 6329 CCGCTGAGCAGCATCAAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGACAGGAC 6388  
Db 7859 CCGCTGAGCAGCATCAAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGACAGGAC 7918  
QY 6389 TAATAAGATACAGCGGTTTCCCGTGAAGTCCCTGCGCTCTCTGTTTCGACCC 6448  
Db 7919 TAATAAGATACAGCGGTTTCCCGTGAAGTCCCTGCGCTCTCTGTTTCGACCC 7978  
QY 6449 TCCCGCTTACCGGATACCTGTCCCGCTTTCTCCCTTCGGGAGCGTGGCGCTTCTCAAT 6508  
Db 7979 TCCCGCTTACCGGATACCTGTCCCGCTTTCTCCCTTCGGGAGCGTGGCGCTTCTCAAT 8038  
QY 6509 GCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCAGGTGGGCTGTGTGC 6568  
Db 8039 GCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCAGGTGGGCTGTGTGC 8098  
QY 6569 ACGAACCCCGCTTCCAGCCCGCTGCTGCTTATCCGTAACCTATCGTCTCTGAGTCCA 6628

8099 AGAAACCCCGTTTCAGCCGACCGCTGCGCTTATCCGGTAACATATCGTCTTGAAGTCCA 8158  
6629 ACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAG 6688  
8159 ACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAG 8218  
6689 CGAGTATGTAGCGGTGCTACAGAGTCTTGAAGTGTGTCGCTTAACATACGGCTACACTA 6748  
8219 CGAGTATGTAGCGGTGCTACAGAGTCTTGAAGTGTGTCGCTTAACATACGGCTACACTA 8278  
6749 GAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTG 6808  
8279 GAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTG 8338  
6809 GTAGCTCTTGATTCGGGCAAAACCAACCCGCTGTGTAGCGGTGTTTTTTTGTGTAAGC 6868  
8339 GTAGCTCTTGATTCGGGCAAAACCAACCCGCTGTGTAGCGGTGTTTTTTTGTGTAAGC 8398  
6869 ACCAGATTACGGCAGAAAAAGAGATCTCAGAGATCCCTTGATCTTTCTACGGGT 6928  
8399 AGCAGATTACGGCAGAAAAAGAGATCTCAGAGATCCCTTGATCTTTCTACGGGT 8458  
6929 CTGACGCTCAGTGGAAACCAACTCAGCTTAAGGATTTTGGTCAATGAGATTATCAAAA 8518  
8459 CTGACGCTCAGTGGAAACCAACTCAGCTTAAGGATTTTGGTCAATGAGATTATCAAAA 8518  
6989 GGATCTTACCTAGATTCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATAT 7048  
8519 GGATCTTACCTAGATTCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATAT 8578  
7049 ATGAGTAACTTGGTCTGACAGTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGA 7108  
8579 ATGAGTAACTTGGTCTGACAGTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGA 8638  
7109 TCTGCTATTTCGTTTATCCATAGTTGCTGACTCCCGTGTGTAGATTAATACTACGATAC 7168  
8639 TCTGCTATTTCGTTTATCCATAGTTGCTGACTCCCGTGTGTAGATTAATACTACGATAC 8698  
7169 GGGGGGCTTACATCTGCCCCAGTGTGCAATCATACCGGACCCAGCCAGCTCACCGG 7228  
8699 GGGGGGCTTACATCTGCCCCAGTGTGCAATCATACCGGACCCAGCCAGCTCACCGG 8758  
7229 CTCACAGATTATCAGCAATAAACCAGCCAGCCGGAAGCCGAGCGAGAGTGTCTCTG 7288  
8759 CTCACAGATTATCAGCAATAAACCAGCCAGCCGGAAGCCGAGCGAGAGTGTCTCTG 8818  
7289 CAATTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTAGAGTAAGTAGTT 7348  
8819 CAATTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTAGAGTAAGTAGTT 8878  
7349 CGCCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTGCAGGCATCTGTTGTCACGCT 7408  
8879 CGCCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTGCAGGCATCTGTTGTCACGCT 8938  
7409 CCGCTGTTGGTATGGCTTCATTCAGTCCGGTTCCCAACGATCAAGGCGAGTTACATGAT 7468  
8939 CCGCTGTTGGTATGGCTTCATTCAGTCCGGTTCCCAACGATCAAGGCGAGTTACATGAT 8998  
7469 CCCCCATGTTGCAAAAAGCGGTAGCTCTCCGTCCTCCGATCGTTGTGCAAGTA 7528  
8999 CCCCCATGTTGCAAAAAGCGGTAGCTCTCCGTCCTCCGATCGTTGTGCAAGTA 9058  
7529 AGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCAGTGCATTAATCTCTTACTGTCA 7588  
9059 AGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCAGTGCATTAATCTCTTACTGTCA 9118  
7589 TGGCATCCGTAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGTCAATTCAGAT 7648  
9119 TGGCATCCGTAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGTCAATTCAGAT 9178  
7649 AGTGTATGGCGGACCGAGTGTCTTGTGCGCGGCTCAACCGGATTAATACCGGCCAC 7708  
9179 AGTGTATGGCGGACCGAGTGTCTTGTGCGCGGCTCAACCGGATTAATACCGGCCAC 9238

7709 ATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTCTTTTCGGGGCGAAAACTCTCAA 7768  
9239 ATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTCTTTTCGGGGCGAAAACTCTCAA 9298  
7769 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTGTGACCCCACTGATCTT 7828  
9299 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTGTGACCCCACTGATCTT 9358  
7829 CAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGCCG 7888  
9359 CAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGCCG 9418  
7889 CAAAAAGGATTAAGGGGACACGGAATTTGTAATCTACTACTCTCTTTTCTTTTCAAT 7948  
9419 CAAAAAGGATTAAGGGGACACGGAATTTGTAATCTACTACTCTCTTTTCTTTTCAAT 9478  
7949 ATTATTGAACATTTATCAGGGTTATTGCTCATAGCGGATACATATTTGAATGTATTT 8008  
9479 ATTATTGAACATTTATCAGGGTTATTGCTCATAGCGGATACATATTTGAATGTATTT 9538  
8009 AGAAAAATAAATAATAGGGTTCCCGGCACATTTCCCGAAAAAGTGCACCTGACGTCT 8068  
9539 AGAAAAATAAATAATAGGGTTCCCGGCACATTTCCCGAAAAAGTGCACCTGACGTCT 9598  
8069 AAGAAACCATTTATCATGACATTAACCTATAAAATAGGCGTATACAGAGCCCTTTTC 8128  
9599 AAGAAACCATTTATCATGACATTAACCTATAAAATAGGCGTATACAGAGCCCTTTTC 9658  
8129 GTCTTCAAGAAATTAATCTCATGTTTGCAGCTTTATCATGATAAGCTGATCATGTTGG 8188  
9659 GTCTTCAAGAAATTAATCTCATGTTTGCAGCTTTATCATGATAAGCTGATCATGTTGG 9718  
8189 TATTGTGAATAGACCGAGTCCGGACACTGGAATAACAGTATTATTTCG 8241  
9719 TATTGTGAATAGACCGAGTCCGGACACTGGAATAACAGTATTATTTCG 9771

## RESULT 10

AX417106

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

i. 9776

/organism="unidentified"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32644"

/note="Plasmid P55"

ORIGIN

Query Match

Best Local Similarity

Matches 8101; Conservative

77.5%; Score 6383.8; DB 6; Length 9776;

Pred. No. 0;

Mismatches 137; Indels 1535; Gaps 10;

2 GATCTAATCCAAAGACGAAAGTGTGAATGAACCTTTTGGCATCCGACATCCACAGG 61

1 GATCTAATCCAAAGACGAAAGTGTGAATGAACCTTTTGGCATCCGACATCCACAGG 60

62 TCCATTCTCACATAAGTCCCAACGCAACAGAGGGGATACACTAGACGACGACCGTT 121

61 TCCATTCTCACATAAGTCCCAACGCAACAGAGGGGATACACTAGACGACGACCGTT 120



Db 2281 TGGCGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATCGCCACCCGTTCTCGGAGC 2340  
Qy 2050 ACTGTCCGACCGCTTTGGCCGCGCCAGTCTCTGCTCGCTTCTGCTTCTGAGCCACTAT 2109  
Db 2341 ACTGTCCGACCGCTTTGGCCGCGCCAGTCTCTGCTCGCTTCTGCTTCTGAGCCACTAT 2400  
Qy 2110 CGACTACCGGATCATGGCGACACACCCGCTCTGCGATCTATCGAATCTAAATGTAAAT 2169  
Db 2401 CGACTACCGGATCATGGCGACACACCCGCTCTGCGATCTATCGAATCTAAATGTAAAT 2460  
Qy 2170 TAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2229  
Db 2461 TAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2520  
Qy 2230 GGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCCAATATGTTTCA 2289  
Db 2521 GGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCCAATATGTTTCA 2580  
Qy 2290 GTCCCTCAGGAGTTACGTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2349  
Db 2581 GTCCCTCAGGAGTTACGTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2640  
Qy 2350 CGCTGTATTGACGGGCATATCGTACGTTGGCAAGTGTGTTGGTACCGGAGGATTAAT 2409  
Db 2641 CGCTGTATTGACGGGCATATCGTACGTTGGCAAGTGTGTTGGTACCGGAGGATTAAT 2700  
Qy 2410 TTCCAACATCTCTGAGAGTAGGACCAACAAACACAGATCCAGCGGTGTTGTTGAT 2469  
Db 2701 CTCCACAACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGCGGTGTTGTTGAT 2760  
Qy 2470 CAACATAGAGAGCAATCTCGATTTTCAGAGTCAAGTGTTCAGGAGCGTACCTGATTTG 2529  
Db 2761 CAACATAGAGAGCAATCTCGATTTTCAGAGTCAAGTGTTCAGGAGCGTACCTGATTTG 2820  
Qy 2530 ACATTTCAAAGCCGTCTGATGTTGCAACCGATAGGTTGTAGAGTGTGCAATACACT 2589  
Db 2821 ACATTTCAAAGCCGTCTGATGTTGCAACCGATAGGTTGTAGAGTGTGCAATACACT 2880  
Qy 2590 TCGGTACATTTCAACCTTGGCACTGACAGCTTGGTGTGACAGCATCTTCAATTC 2649  
Db 2881 TCGGTACATTTCAACCTTGGCACTGACAGCTTGGTGTGACAGCATCTTCAATTC 2940  
Qy 2650 TGGCAAGCTCTGTTCTGATATCGACAGCCCAACAGATCACTGGGAATCAATACCAT 2709  
Db 2941 TGGCAAGCTCTGTTCTGATATCGACAGCCCAACAGATCACTGGGAATCAATACCAT 3000  
Qy 2710 GTTCAGTTGAGACAGAGGTTGAGGCAACGAATCTGGATCGAGCTATTTATCAGCAA 2769  
Db 3001 GTTCAGTTGAGACAGAGGTTGAGGCAACGAATCTGGATCGAGCTATTTATCAGCAA 3059  
Qy 2770 TAACTAGAACTTCAGAGCCCGCAGCAGCATGTCAATCTACTACAGAGGCTCATGTGTCAT 2829  
Db 3060 TAACTAGAACTTCAGAGCCCGCAGCAGCATGTCAATCTACTACAGAGGCTCATGTGTCAT 3119  
Qy 2830 TTTGAACCATCATCTTGGCAGCAGTAAACAACTGTTTCTGGACCAATATTTGTGTCAC 2889  
Db 3120 TTTGAACCATCATCTTGGCAGCAGTAAACAACTGTTTCTGGACCAATATTTGTGTCAC 3179  
Qy 2890 ACTTAGGACAGTTTCTGTTCCGTAAGCATAGAGCTACTGCTGGGGCGCTCTGCTA 2949  
Db 3180 ACTTAGGACAGTTTCTGTTCCGTAAGCATAGAGCTACTGCTGGGGCGCTCTGCTA 3239  
Qy 2950 GCACGATACACTTAGCACCACCTTGTGGGCAACGATAGATGACTTCTGGGGTAAGGGTAC 3009  
Db 3240 GCACGATACACTTAGCACCACCTTGTGGGCAACGATAGATGACTTCTGGGGTAAGGGTAC 3299  
Qy 3010 CATCTCTTAGGTGAGTAGGAGGCAAAACAAATTTCTTGGCAACGACCACTTTGGCAGGAA 3069  
Db 3300 CATCTCTTAGGTGAGTAGGAGGCAAAACAAATTTCTTGGCAACGACCACTTTGGCAGGAA 3359  
Qy 3070 CACCAGCATCAGGAGTAGGAGGCAAAATTTGGGTTCCACGAGGATATAGAGGCCAA 3129  
Db 3360 CACCAGCATCAGGAGTAGGAGGCAAAATTTGGGTTCCACGAGGATATAGAGGCCAA 3419

Qy 3130 CTTTCTCAATAGTCTTTGCAAAACAGAGACAGACTACACAGGGCAAGTCTCAACTTGA 3189  
Db 3420 CTTTCTCAATAGTCTTTGCAAAACAGAGACAGACTACACAGGGCAAGTCTCAACTTGA 3479  
Qy 3190 ACCTCTCCGTTAGTGTGAGCTTCAATGAATTTCTGACGTTTCTATAGAGAGTCAATGG 3249  
Db 3480 ACCTCTCCGTTAGTGTGAGCTTCAATGAATTTCTGACGTTTCTATAGAGAGTCAATGG 3539  
Qy 3250 CTCTCTTAAGCTTATCTGCAATTTGCAATTAAGTTCCTCTGGAAAGAGCTTCTAAACACAG 3309  
Db 3540 CTCTCTTAAGCTTATCTGCAATTTGCAATTAAGTTCCTCTGGAAAGAGCTTCTAAACACAG 3599  
Qy 3310 GTCTCTTAAAGCACTCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTTTGTACCAAT 3369  
Db 3600 GTCTCTTAAAGCACTCCATCAAACTTGGCAGTTAGTTCTAAAGGGCTTTGTACCAAT 3659  
Qy 3370 TTTGACGACATTTGTCGCAATTTGTCATTAATTTCCATATCTCTTCGTTTCTGGA 3429  
Db 3660 TTTGACGACATTTGTCGCAATTTGTCATTAATTTCCATATCTCTTCGTTTCTGGA 3719  
Qy 3430 TAGGACGACGAAGGGCATCTTCAATTTCTTGTGAGAGGCTTTAGAAACGTCATTTTGC 3489  
Db 3720 TAGGACGACGAAGGGCATCTTCAATTTCTTGTGAGAGGCTTTAGAAACGTCATTTTGC 3779  
Qy 3490 ACAATTTCAATACGACCTTCAGAGGACCTTCTTAGTTTGGATTTCTTCTTAGGTTGT 3549  
Db 3780 ACAATTTCAATACGACCTTCAGAGGACCTTCTTAGTTTGGATTTCTTCTTAGGTTGT 3839  
Qy 3550 CTTTGTGTATCTCTGCTTGGCATCTCTTCTCTAGTGACCTTTAGGACCTTCATAT 3609  
Db 3840 CTTTGTGTATCTCTGCTTGGCATCTCTTCTCTAGTGACCTTTAGGACCTTCATAT 3899  
Qy 3610 CCAGGTTTCTCTCCACTCGTCCAACTGACACCTGACCTTTGGCACATCTAACTAATGCAA 3669  
Db 3900 CCAGGTTTCTCTCCACTCGTCCAACTGACACCTGACCTTTGGCACATCTAACTAATGCAA 3959  
Qy 3670 AATAAATAAGTACGACATCTCCAGGCTATATCTTCTTGGATTTAGCTTCTGCAAGTT 3729  
Db 3960 AATAAATAAGTACGACATCTCCAGGCTATATCTTCTTGGATTTAGCTTCTGCAAGTT 4019  
Qy 3730 CATCAGCTTCTCTCCACTTTAGCGTTC-AAACAAACTTCGTCGCAAACTAACCGTTG 3788  
Db 4020 CATCAGCTTCTCTCCACTTTAGCGTTC-AAACAAACTTCGTCGCAAACTAACCGTTG 4079  
Qy 3789 GTATAAGAACTTCTGAGGANTGCTTCTTACGATCCCAAGTGCTTCCATGGCTCTA 3848  
Db 4080 GTATAAGAACTTCTGAGGANTGCTTCTTACGATCCCAAGTGCTTCCATGGCTCTA 4138  
Qy 3849 AGACCTTTGATTTGGCCAAAACAGGAAGTCCGTTTCCAACTGACAGAAACCAACCTGTT 3908  
Db 4139 AGACCTTTGATTTGGCCAAAACAGGAAGTCCGTTTCCAACTGACAGAAACCAACCTGTT 4198  
Qy 3909 TGTTCACCAAAATTTCAAGCAGTCTCCATCAATCCAAATTCGATCCAGCAACTTT 3968  
Db 4199 TGTTCACCAAAATTTCAAGCAGTCTCCATCAATCCAAATTCGATCCAGCAACTTT 4258  
Qy 3969 TGAGTTGCTTCCAGATCTGACACTTTTATACCAAACTGACAGAGATTTGGTAGACT 4028  
Db 4259 TGAGTTGCTTCCAGATCTGACACTTTTATACCAAACTGACAGAGATTTGGTAGACT 4318  
Qy 4029 CCAGTTTGTCTCTTATAGCCTCCGAAATAGACTTTTGGACGAGTACACAGGCCAAC 4088  
Db 4319 CCAGTTTGTCTCTTATAGCCTCCGAAATAGACTTTTGGACGAGTACACAGGCCAAC 4378  
Qy 4089 GAGTAAATAGAGAGTCAGCCCAAGTAGTGAATAGACCATTCGGGGCGGTGAGTGC 4148  
Db 4379 GAGTAAATAGAGAGTCAGCCCAAGTAGTGAATAGACCATTCGGGGCGGTGAGTGC 4438  
Qy 4149 AAAGAGGCCCAACAAAATTTCACTGACAGGGAATTTTTCACATCTTTCAGAAAGTTCGTAT 4208  
Db 4439 AAAGAGGCCCAACAAAATTTCACTGACAGGGAATTTTTCACATCTTTCAGAAAGTTCGTAT 4498

QY 4209 TCAGTAGTCAATTCGCCGAGCATCAATAATGGGATATATACAGAAGCAACAGTGGAGTC 4268  
Db 4499 TCAGTAGTCAATTCGCCGAGCATCAATAATGGGATATATACAGAAGCAACAGTGGAGTC 4558  
QY 4269 ACATCTACCAACTTTGGGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGCCATTAGTG 4328  
Db 4559 ACATCTACCAACTTTGGGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGCCATTAGTG 4618  
QY 4329 AAACTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCATCTAGCATTAGCGGC 4388  
Db 4619 AAACTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCATCTAGCATTAGCGGC 4678  
QY 4389 AAGGATGCAACTTTATCAACACAGGTCCTATAGATAAACCTAGCGCTGGGATCATCCTT 4448  
Db 4679 AAGGATGCAACTTTATCAACACAGGTCCTATAGATAAACCTAGCGCTGGGATCATCCTT 4738  
QY 4449 TGGACAACTCTTCTGCCAATCTAGTCCAAATCACTTCAATGTATACCATTTATGTAC 4508  
Db 4739 TGGACAACTCTTCTGCCAATCTAGTCCAAATCACTTCAATGTATACCATTTATGTAC 4798  
QY 4509 AACTTGACAAAGTTGTGATCAGCTCCTCAAAATGGTCTCTGTAAACGGATGACTCAACT 4568  
Db 4799 AACTTGACAAAGTTGTGATCAGCTCCTCAAAATGGTCTCTGTAAACGGATGACTCAACT 4858  
QY 4569 TGCACATTAATCTGAAGTCTAGTCTGATGAGTGAACCTGATCAGGTTGTGCAGTGGTCA 4628  
Db 4859 TGCACATTAATCTGAAGTCTAGTCTGATGAGTGAACCTGATCAGGTTGTGCAGTGGTCA 4918  
QY 4629 GCAGATAGGAAACACCGCTTTTCTTACCAAACTCAAGGAATATCAAACTCTGCAACA 4688  
Db 4919 GCAGATAGGAAACACCGCTTTTCTTACCAAACTCAAGGAATATCAAACTCTGCAACA 4978  
QY 4689 CTTGCGTATGAGGTAGCAAGGAAATGTCTATCTGAAGTCGACAGTGAAGTGTAGTCT 4748  
Db 4979 CTTGCGTATGAGGTAGCAAGGAAATGTCTATCTGAAGTCGACAGTGAAGTGTAGTCT 5038  
QY 4749 TGAGAAATCTGAGCCGATTTTATTTATTCAGTGAAGTCACTATCAGGAGATCTCTAC 4808  
Db 5039 TGAGAAATCTGAGCCGATTTTATTTATTCAGTGAAGTCACTATCAGGAGATCTCTAC 5098  
QY 4809 GCCGACGATCTGGCCGACCTGCAGGTCCGCATACCGCCGCCACAGGTGCGTGTCT 4868  
Db 5099 GCCGACGATCTGGCCGACCTGCAGGTCCGCATACCGCCGCCACAGGTGCGTGTCT 5158  
QY 4869 GCGCCTATATCGCGACATCAACGATGGGAAGATCGGCTCGCCACTTCGGGCTCATG 4928  
Db 5159 GCGCCTATATCGCGACATCAACGATGGGAAGATCGGCTCGCCACTTCGGGCTCATG 5218  
QY 4929 AGCGCTTGTTCGGGCTGGTATGTTGTCAGGCGCCCGTGGCCGGGGGACTGTTGGGCGCC 4988  
Db 5219 AGCGCTTGTTCGGGCTGGTATGTTGTCAGGCGCCCGTGGCCGGGGGACTGTTGGGCGCC 5278  
QY 4989 ATCTCCTTG----- 4997  
Db 5279 ATCTCCTTGACCTGCAGGGGGGGGGGAAAGCCACGTTGTGTCTCAAAATCTGTAT 5338  
QY 4998 ----- 4997  
Db 5339 GTTACATTGCACAGATAAATAATATATCATCATGACAAATAAATCTGCTTACATAA 5398  
QY 4998 ----- 4997  
Db 5399 ACAGTAATACAGGGGTGTATGAGCCATATTCAACGGGAAACGCTCTTGTCTAAGGCGCC 5458  
QY 4998 ----- 4997  
Db 5459 GATTAAATCCAAATGATGCTGATTTATATGGGTATTAATGGGCTCGCATTAATGTGCG 5518  
QY 4998 ----- 4997  
Db 5519 GGCAATCAGGTGCGACAAATCTATCATGTTGTATGGGAAGCCGATGCGCGAGTGTTC 5578  
QY 4998 ----- 4997

Db 5579 TGAACATGCGAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGTCAGACTAAACT 5638  
QY 4998 ----- 4997  
Db 5639 GGCTGACGGAATTTATGCTCTCTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATG 5698  
QY 4998 ----- 4997  
Db 5699 CATGTTACTCACCACCTGCGATCCCGGAAAAACAGCATTCACAGGTATTAGAAATATC 5758  
QY 4998 ----- 4997  
Db 5759 CTGATTGAGTGAAAAATATTGTTGATGCGCTGGCAGTGTCTCTCGCGGTTGCAATCGA 5818  
QY 4998 ----- 4997  
Db 5819 TTCTCTGTTTGAATTTCTCTTTTAAACAGCGATCGGTATTTCGTCTCGCTCAGCGCAAT 5878  
QY 4998 ----- 4997  
Db 5879 CACGAATGAATAACGTTTGGTTGATGCGGAGTGATTTTGATGACGAGCGTAATAGTGTG 5938  
QY 4998 ----- 4997  
Db 5939 CTGTTGAACAGTCTGGAAGAAATGCATAAGCTTTTCCCATTTCTCACCGGATTCACTCG 5998  
QY 4998 ----- 4997  
Db 5999 TCACCTCATGGTATTCTCACTTGATNAACCTTATTTTGAAGGGAATTAATAGTGT 6058  
QY 4998 ----- 4997  
Db 6059 GTATTGATGTTGACAGCTCGGAATCGCAGACCGATACAGGATCTTGCCCATCCTATGGA 6118  
QY 4998 ----- 4997  
Db 6119 ACTGCTCGGTGAGTTTCTCTCTCAATTAACAGAAACGCTTTTTCAAAAATATGTTATG 6178  
QY 4998 ----- 4997  
Db 6179 ATAATCTGATATGAATAAATGTCAGTTTCAATTCATCTCGATGAGTTTCTTAATCAG 6238  
QY 4998 ----- 4997  
Db 6239 AATGTTAATTTGGTTGTAACACTGGCAGAGCATTAACGCTGACGTGACGGGACGCGGCT 6298  
QY 4998 ----- 4997  
Db 6299 TTGTTGAATAAATCGAACTTTTGTGCTGAGTGAAGGATCAGATCAGCATCTTCCGACAA 6358  
QY 4998 ----- 4997  
Db 6359 CGCAGACCGTTCCGTTGGCAAGAAAAAGTTCAAAATCACCAACTGGTCCACCTACAACAA 6418  
QY 4998 ----- 4997  
Db 6419 AGCTCTCATFAACCGTGGCTCCCTCACTTTCTGGCTGATGATGGGGGATTCAGGCCGTG 6478  
QY 4998 ----- 5008  
Db 6479 GTATGATGACCAACACCTTCTTTCACAGGACAGACCTCAGCGCCCCCCCCCTGCAGG 6538  
QY 5009 CTTTGGCGGGCGGTCTCAACGGCTCAACCTACTACTGGGCTGCTTCTTCTTAATGACAGA 5068  
Db 6539 TCCACGCGGGCGGTCTCAACGGCTCAACCTACTACTGGGCTGCTTCTTCTTAATGACAGA 6598  
QY 5069 GTGCGATAAGGAGAGCGTCGAGTATCTATGATTTGGAAGTATGGGAATGGTATACCCGC 5128  
Db 6599 GTGCGATAAGGAGAGCGTCGAGTATCTATGATTTGGAAGTATGGGAATGGTATACCCGC 6658  
QY 5129 ATTCTTCAGTGTGAGGTCTCCTATCAGTATGCCCCAACTAAAGCAACCGGAGGAGG 5188

Db 6659 ATTCTTCAGTGTCTTGGTCTCTCTATCAGAAATGAGCCCAACTAAACCCGAGGAGG 6718  
Qy 5189 AGATTTTCATGTAATTTCTGACTTTTGGTCTATCAGTAGACTCGAACTGTGAGACTAT 5248  
Db 6719 AGATTTTCATGTAATTTCTGACTTTTGGTCTATCAGTAGACTCGAACTGTGAGACTAT 6778  
Qy 5249 CTCGGTTATGACAGAGAAATGTCTCTTGGAGACAGTAATGAATGCCACCAATAAA 5308  
Db 6779 CTCGGTTATGACAGAGAAATGTCTCTTGGAGACAGTAATGAATGCCACCAATAAA 6838  
Qy 5309 GAAATCCTTTGTTATCAGGAACAAACTCTCTGTTTGGAACTTTTTCGGTGCCTTGAACATAT 5368  
Db 6839 GAAATCCTTTGTTATCAGGAACAAACTCTCTGTTTGGAACTTTTTCGGTGCCTTGAACATAT 6898  
Qy 5369 AAAATGTAGAGTGATATGTCCGGTAGGAATGGAGCGGCAAAATGCTTACCTTCGGACC 5428  
Db 6899 AAAATGTAGAGTGATATGTCCGGTAGGAATGGAGCGGCAAAATGCTTACCTTCGGACC 6958  
Qy 5429 TTCAAGAGGTATGAGGTTTGTAGATCTATGCAACTTTCAGTGACAAAGTGTCTATT 5488  
Db 6959 TTCAAGAGGTATGAGGTTTGTAGATCTATGCAACTTTCAGTGACAAAGTGTCTATT 7018  
Qy 5489 TCGTTCAAACCAATCCGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATTGATCCAAAGC 5548  
Db 7019 TCGTTCAAACCAATCCGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATTGATCCAAAGC 7078  
Qy 5549 CAGTCGGTCTTGAACACTGACAAATAGTGTCTGCTGTTTTCAGGTCTATCTTTGATGAAT 5608  
Db 7079 CAGTCGGTCTTGAACACTGACAAATAGTGTCTGCTGTTTTCAGGTCTATCTTTGATGAAT 7138  
Qy 5609 AAATCTAGTCTTTGATCTAAATTAATCTTGAAGCGGCAAGGATAAATACCAAAATCTAA 5668  
Db 7139 AAATCTAGTCTTTGATCTAAATTAATCTTGAAGCGGCAAGGATAAATACCAAAATCTAA 7198  
Qy 5669 AACTCTTTTAAACGTTTAAAGGACAAATATGTCTGCTGTATTAACCCCAAAATCTAGCT 5728  
Db 7199 AACTCTTTTAAACGTTTAAAGGACAAATATGTCTGCTGTATTAACCCCAAAATCTAGCT 7258  
Qy 5729 CGTAGCTGATCTCATCACTGAGGAGCACTATCTTGTGTTTGAAGAAATTCGGGAGA 5788  
Db 7259 CGTAGCTGATCTCATCACTGAGGAGCACTATCTTGTGTTTGAAGAAATTCGGGAGA 7318  
Qy 5789 TCGGATATCGAGAAAGGATCGCTGATTTTAAACGTAATTTATCTCAAGATCTCTGC 5848  
Db 7319 TCGGATATCGAGAAAGGATCGCTGATTTTAAACGTAATTTATCTCAAGATCTCTGC 7378  
Qy 5849 CTCGCGGTTTCGGTGATGACGGTGAAACCTCTGACACATGCGACTCCCGGAGACGTC 5908  
Db 7379 CTCGCGGTTTCGGTGATGACGGTGAAACCTCTGACACATGCGACTCCCGGAGACGTC 7438  
Qy 5909 ACAGCTTGTCTGTAAGCGGATCCCGGAGCAGACAGCCCGTCAAGGCGCGTCAAGCGGT 5968  
Db 7439 ACAGCTTGTCTGTAAGCGGATCCCGGAGCAGACAGCCCGTCAAGGCGCGTCAAGCGGT 7498  
Qy 5969 GTTGGCGGTGTCCGGGCGACGCAATGACCAAGTCAAGTACGAGTACGAGTGTATCT 6028  
Db 7499 GTTGGCGGTGTCCGGGCGACGCAATGACCAAGTCAAGTACGAGTGTATCT 7558  
Qy 6029 GCTTTAACTATCGGCATCAGAGCAGATTGTACTGAGAGTGCAACCATATCGGTTGAAA 6088  
Db 7559 GCTTTAACTATCGGCATCAGAGCAGATTGTACTGAGAGTGCAACCATATCGGTTGAAA 7618  
Qy 6089 TACCGCACAGATCGCTAAGAGAAATACCGCATCAGGCGCTCTTCCGTTCTCTGCTCA 6148  
Db 7619 TACCGCACAGATCGCTAAGAGAAATACCGCATCAGGCGCTCTTCCGTTCTCTGCTCA 7678  
Qy 6149 CTGACTCGCTGCGCTCGGTCTGCTGCGGAGCGGTATCAGTCACTCAAAAGCGG 6208  
Db 7679 CTGACTCGCTGCGCTCGGTCTGCTGCGGAGCGGTATCAGTCACTCAAAAGCGG 7738  
Qy 6209 TAATACGGTTATCCACAGAAATCAGGGGATAAGCAGAGAAAGCAATGTGAGCAAAAGGCC 6268  
Db 7739 TAATACGGTTATCCACAGAAATCAGGGGATAAGCAGAGAAAGCAATGTGAGCAAAAGGCC 7798

Qy 6269 AGCAAAAGGCCAGGAAACCGTAAAAAGGCGCGTTCCTGGCGTTTTCATAGGCTCCGCC 6328  
Db 7799 AGCAAAAGGCCAGGAAACCGTAAAAAGGCGCGTTCCTGGCGTTTTCATAGGCTCCGCC 7858  
Qy 6329 CCCTGACAGCATCACAAAATTCAGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGAC 6388  
Db 7859 CCCTGACAGCATCACAAAATTCAGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGAC 7918  
Qy 6389 TATAAAGATACAGAGGCTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCC 6448  
Db 7919 TATAAAGATACAGAGGCTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCC 7978  
Qy 6449 TCGCGCTTACCGGATACCTGTCGCGCTTTCCTTCGGAAGCGTGGCGCTTCTCAT 6508  
Db 7979 TCGCGCTTACCGGATACCTGTCGCGCTTTCCTTCGGAAGCGTGGCGCTTCTCAT 8038  
Qy 6509 GTTCACGCTGTAGGTATCTCAGTTCGGTGTAGTCTGTTCCCTCAAGCTGGGCTGTGTC 6568  
Db 8039 GTTCACGCTGTAGGTATCTCAGTTCGGTGTAGTCTGTTCCCTCAAGCTGGGCTGTGTC 8098  
Qy 6569 ACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCTTAACCTATCTGTTGAGTCCA 6628  
Db 8099 ACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCTTAACCTATCTGTTGAGTCCA 8158  
Qy 6629 ACCCGTAAGACACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG 6688  
Db 8159 ACCCGTAAGACACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG 8218  
Qy 6689 CGAGGTATGTAGGCGGTGTACAGATCTTGAAGTGGTGGCTTAACCTAGCGGTACACTA 6748  
Db 8219 CGAGGTATGTAGGCGGTGTACAGATCTTGAAGTGGTGGCTTAACCTAGCGGTACACTA 8278  
Qy 6749 GAAGGACAGTATTTGTTGATCTGCTGAGCGCAGTTACTTCGGAAAAAGAGTTG 6808  
Db 8279 GAAGGACAGTATTTGTTGATCTGCTGAGCGCAGTTACTTCGGAAAAAGAGTTG 8338  
Qy 6809 GTAGCTCTTGATCCGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTTTTGTGCAAGC 6868  
Db 8339 GTAGCTCTTGATCCGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTTTTGTGCAAGC 8398  
Qy 6869 AGCAGATTACCGCAGAAAAAGGATCTCAGAGATCCTTTGATCTTTCTACGCGGT 6928  
Db 8399 AGCAGATTACCGCAGAAAAAGGATCTCAGAGATCCTTTGATCTTTCTACGCGGT 8458  
Qy 6929 CTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGGTCTATGAGATTATCAAAA 6988  
Db 8459 CTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGGTCTATGAGATTATCAAAA 8518  
Qy 6989 GGATCTTTACCTAGATCTCTTTTAAATTAAGATTAAGTTTAAATCAATCTAAAGTATAT 7048  
Db 8519 GGATCTTTACCTAGATCTCTTTTAAATTAAGATTAAGTTTAAATCAATCTAAAGTATAT 8578  
Qy 7049 ATGAGTAAACTTCGCTGACAGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGA 7108  
Db 8579 ATGAGTAAACTTCGCTGACAGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGA 8638  
Qy 7109 TCTGCTTATTTCTGTTCAATAGTTGCTGATCTCCCGCTCGTGTAGATAACTACGATAC 7168  
Db 8639 TCTGCTTATTTCTGTTCAATAGTTGCTGATCTCCCGCTCGTGTAGATAACTACGATAC 8698  
Qy 7169 GGGAGGCTTACCATCTGCGCCAGTCTGCAATGATACCGGAGACCCGACGCTCACCGG 7228  
Db 8699 GGGAGGCTTACCATCTGCGCCAGTCTGCAATGATACCGGAGACCCGACGCTCACCGG 8758  
Qy 7229 CTCAGATTTTATCAGCAATAAACACAGCAGCCGGAAGGCGGAGCGAGAGTGGTCTG 7288  
Db 8759 CTCAGATTTTATCAGCAATAAACACAGCAGCCGGAAGGCGGAGCGAGAGTGGTCTG 8818  
Qy 7289 CAATTTTATCGGCTCCATCAGTCTATTAATTTGTCGCGGAGCTAGAGTAGTAGTT 7348  
Db 8819 CAATTTTATCGGCTCCATCAGTCTATTAATTTGTCGCGGAGCTAGAGTAGTAGTT 8878



7349 GCCAGTTAAAGTGGTGGCAACGCTGTTGGTCATTGCTGAGGATCGTGGTGTGACGCT 7408  
Db GCCAGTTAAAGTGGTGGCAACGCTGTTGGTCATTGCTGAGGATCGTGGTGTGACGCT 8338  
7409 CGTGGTGGTATGCTTCAATCAGTCCCGTTCCTCCAAACGATCAAGCGGAGTTACATGAT 7468  
Db CGTGGTGGTATGCTTCAATCAGTCCCGTTCCTCCAAACGATCAAGCGGAGTTACATGAT 8998  
7469 CCCCATGCTGTGCAAAAAGCGTTAGCTCTCTTGGTCTCCATCGTGTGAGAGTA 7528  
Db CCCCATGCTGTGCAAAAAGCGTTAGCTCTCTTGGTCTCCATCGTGTGAGAGTA 9058  
7529 AGTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGTGCATATCTCTTACTGTCA 7588  
Db AGTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGTGCATATCTCTTACTGTCA 9118  
7589 TGCCATCCGTAAAGTCTTTCTCTGACTGTGAGTACTCAACCAAGTCAATCTGAGAT 7648  
Db TGCCATCCGTAAAGTCTTTCTCTGACTGTGAGTACTCAACCAAGTCAATCTGAGAT 9178  
7649 AGTGTATCGCGCGACCGAGTTGCTCTTGGCCGCGTCAACACGGGATATACCGCGCCAC 7708  
Db AGTGTATCGCGCGACCGAGTTGCTCTTGGCCGCGTCAACACGGGATATACCGCGCCAC 9238  
7709 ATAGCAGAACTTTAAAGTGTCTCATGTTGGAAGAAAGTCTTCCGGCGGAAACTCTCAA 7768  
Db ATAGCAGAACTTTAAAGTGTCTCATGTTGGAAGAAAGTCTTCCGGCGGAAACTCTCAA 9298  
7769 GGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGACCACTGATCTT 7828  
Db GGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGACCACTGATCTT 9358  
7829 CAGCATCTTTTACTTACCAGCGTTCCTGGTGAGCAAAAACAGGAAGGCAAAATCCG 7888  
Db CAGCATCTTTTACTTACCAGCGTTCCTGGTGAGCAAAAACAGGAAGGCAAAATCCG 9418  
7889 CAAAAAGGGAATAGGCGCAGCAAGAAATGCTCAATCACTACTCTCTCTTTTCAAT 7948  
Db CAAAAAGGGAATAGGCGCAGCAAGAAATGCTCAATCACTACTCTCTCTTTTCAAT 9478  
7949 ATTATGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTCAATGATTT 8008  
Db ATTATGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTCAATGATTT 9538  
8009 AGAAAAATAAACAATAGGGTTCGGCGCAGATTTCCCGAAGTCCACCTGACGCT 8068  
Db AGAAAAATAAACAATAGGGTTCGGCGCAGATTTCCCGAAGTCCACCTGACGCT 9598  
8069 AAGAAACCATTTATCATGATTAACCTATAAATAAGCGGTATCAGGAGCCCTTTC 8128  
Db AAGAAACCATTTATCATGATTAACCTATAAATAAGCGGTATCAGGAGCCCTTTC 9598  
8129 GTCTTCAAGATTAATCTCTGATTTGACGCTTATCATGATGAGTCAATGTTG 8188  
Db GTCTTCAAGATTAATCTCTGATTTGACGCTTATCATGATGAGTCAATGTTG 9718  
8189 TATTGTGAATAGCGGATCGGAAACACTGAAATAATACAGTTATTATTCG 8241  
Db TATTGTGAATAGCGGATCGGAAACACTGAAATAATACAGTTATTATTCG 9771

RESULT 11  
LOCUS AY178634 9139 bp DNA circular SYN 06-APR-2003  
DEFINITION Expression vector pYPX99, complete sequence.  
ACCESSION AY178634  
VERSION AY178634.1 GI:29568880  
KEYWORDS Expression vector pYPX99  
ORGANISM Expression vector pYPX99  
SOURCE artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 9139)  
AUTHORS Yao, Q., Peng, R. and Xiong, A.

TITLE Direct Submission  
JOURNAL Submitted (12-NOV-2002) Shanghai Academy of Agricultural Sciences,  
Agro-Biotechnology Research Center, 2901 Beidi Rd, Shanghai 201106,  
China  
COMMENT NCBI staff are still waiting for submitters to provide appropriate  
coding region information.  
FEATURES  
Location/Qualifiers  
source  
1..9139  
/organism="Expression vector pYPX99"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:218468"  
/lab\_host="Pichia pastoris"  
misc\_feature  
1..936  
/note="modified AOX1 promoter"  
misc\_feature  
937..1212  
/note="similar to modified alpha factor secretion signal"  
terminator  
1275..1607  
/note="AOX1 transcription terminator"  
misc\_feature  
complement(1994..4527)  
/note="similar to phosphoribosyl-ATP pyrophosphohydrolase"  
CDS  
4798..5613  
/codon\_start=1  
/product="aminoglycoside-3'-O-phosphotransferase"  
/protein\_id="AA084033.1"  
/db\_xref="GI:29568881"  
/translation="MSHQRETSCSRPLNSNMDADLYGKWARDNVQSGATVRLY  
GKPAPELFLKHGSGVANDVTDEMRLNMLTEFMPFTIKHFIPTDDAWLLTALP  
KTFAPVLEEDYDGSNIVDALVFLRLHSIPVNCFFNSDRVFLRAQASNNVGL  
VDASDFDERNGWPEQVQKEMHKLFPSPDSVTHGDFSLDNLIFDEGLIGCIDVG  
RVGIADRYQDLAILWNLCEGPEPSLQKLFQKYGIDNPDMNKLOFHLMLDEFF"  
5987..6736  
/note="3' AOX1 fragment from Pichia pastoris"  
misc\_feature  
complement(7151..7824)  
/note="ColE1"  
rep\_origin  
complement(7969..8829)  
/function="ampicillin resistance"  
CDS  
/codon\_start=1  
/product="beta-lactamase"  
/protein\_id="AA084034.1"  
/db\_xref="GI:29568882"  
/translation="MSIQHFRVALIPFAARCLPVFAHPETLVKVKDAEDOLGARVGY  
IELDSKILLESFRRPEPMSTFKVLGAVLSVDAGQEQGLRRIRHYSDNLVE  
YSPVTERLDTGMTVRELCSAAITMSDNTAANLLLTIGGPKELTALHNHGDVRL  
DRWEPELNEAIPNDERDTPAAMATTLIRKLLTGLLELLTASRQQLIDWMDKRVAGPL  
LRSLALPWGFIADKSGAGERSGLIAALGPDGKPSRIVTVYTTGSOATMDERNQIA  
EIGASLKHWH"

## ORIGIN

Query Match 75.8%; Score 6250.2; DB 12; Length 9139;  
Best Local Similarity 83.7%; Pred. No. 0;  
Matches 7817; Conservative 0; Mismatches 228; Indels 1290; Gaps 14;  
QY 1 AGATCTACATCCCAAGCAGAAAGGTTCAATGAACACCTTTTGGCATCCGACATCCACAG 60  
Db 1 AGATCTACATCCCAAGCAGAAAGGTTCAATGAACACCTTTTGGCATCCGACATCCACAG 60  
QY 61 GTCCATTCTCACATAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGACCGT 120  
Db 61 GTCCATTCTCACATAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGACCGT 120  
QY 121 TCGAAACGAGGACCTCCACTCTTCTCTCTACACCGCCTTTTGGCATCGAAACACC 180  
Db 121 TCGAAACGAGGACCTCCACTCTTCTCTCTACACCGCCTTTTGGCATCGAAACACC 180  
QY 181 AGCCCAAGTTATTGGCTTGAATGGAGCTCGCTCAATTCCTTCTATTAGGCTACTA 240  
Db 181 AGCCCAAGTTATTGGCTTGAATGG---CGCTCAATTCCTTCTATTAGGCTACTA 236  
QY 241 ACACCATGACATTTATAGCTGTCTATCTGGCCCCCTGGCGAGGTTCAATGTTGTTA 300  
Db 237 ACACCATGACATTTATAGCTGTCTATCTGGCCCCCTGGCGAGGTTCAATGTTGTTA 296  
QY 301 TTTCCGAATGCAACAGCTCCGATTACACCCGCAATCACTCCAGATGAGGGCTTCTG 360

Db 297 TTTCGAAATGCAACAAAGCTCGCATTAACCCGAAATCACTCCAGATGAGGCTTTCTG 356  
Qy 361 AGTGTGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAAATGACACAGTTTAAACGCT 420  
Db 357 AGTGTGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAAATGACACAGTTTAAACGCT 416  
Qy 421 GTCTTGGAACTATATATGACAAAAGCGTGATCTCATCCAAAGATGAACATAAGTTTGGTTG 480  
Db 417 GTCTTGGAACTATATATGACAAAAGCGTGATCTCATCCAAAGATGAACATAAGTTTGGTTG 476  
Qy 481 TTGAAATGCTAAACGGCCAGTTGGTCAAAAGAAACTTCCAAAGTCCGCAATACCGTTTGT 540  
Db 477 TTGAAATGCTAAACGGCCAGTTGGTCAAAAGAAACTTCCAAAGTCCGCAATACCGTTTGT 536  
Qy 541 CTGTGTTGGTATTCATTTGACAAATGCTCAAAATATCTCATTAATGCTTTAGCGCAGTCT 600  
Db 537 CTGTGTTGGTATTCATTTGACAAATGCTCAAAATATCTCATTAATGCTTTAGCGCAGTCT 596  
Qy 601 CTCTATCGCTTCTGAACCCCGGTGACCTGTGCGGAACGCAATGGGAAACACCCGCT 660  
Db 597 CTCTATCGCTTCTGAACCCCGGTGACCTGTGCGGAACGCAATGGGAAACACCCGCT 656  
Qy 661 TTTTGGATGATTAATGCAATTTGCTCCACATTTGTATGCTTCCAAGATTTCTGTGGGAATACT 720  
Db 657 TTTTGGATGATTAATGCAATTTGCTCCACATTTGTATGCTTCCAAGATTTCTGTGGGAATACT 716  
Qy 721 GCTGATAGCTAAAGCTTCAATGATCAAAATTTAACTGTCTAACCCCTACTTGACAGCAAT 780  
Db 717 GCTGATAGCTTAAAGCTTCAATGATCAAAATTTAACTGTCTAACCCCTACTTGACAGCAAT 776  
Qy 781 ATATAACAGAAAGGAGCTGCTCTTAACCTTTTATCATCAATTAAGCTT 840  
Db 777 ATATAACAGAAAGGAGCTGCTCTTAACCTTTTATCATCAATTAAGCTT 836  
Qy 841 ACTTTCATPAATGGCACTGGTTCCAAATGACAAGCTTTGATTTAAGCACTTTTAAACGA 900  
Db 837 ACTTTCATPAATGGCACTGGTTCCAAATGACAAGCTTTGATTTAAGCACTTTTAAACGA 896  
Qy 901 CAACCTTGAGAGATCAAAACCACTAATTAATTCGAGGATCCAAAGATGAGATTTCT 960  
Db 897 CAACCTTGAGAGATCAAAACCACTAATTAATTCGAGGATTTCTCAACCA 951  
Qy 961 TCAATTTTACTGCAAGTTTATTCGAGCATCTCCGCAATAGCTGTCTCCAGTCAACACT 1020  
Db 952 TCTATCTTCACTGCTGTCTTGTGCTGCTCTTCTGCTGCTTGGTGTCTCTGTCAACACT 1011  
Qy 1021 ACAACAGAGATGAACCGCAAAATTCGGGTGAGCTGTCACTGGTTACTCTCAGATTTA 1080  
Db 1012 ACTACCGAGGATGAACCTGCTCAAAATCCCTGTGAGGCTGTCACTGGTTACTCTGACCTG 1071  
Qy 1081 GAAGGGGATTTGCAATGCTGTTTGGCAATTTTCCAAACAGACAAATAACGGGTTATTG 1140  
Db 1072 GAGGGTGACTTGAAGCTGCTGTCTTGGCAATTTCTTAATCCACCAACACGGTTTGTG 1131  
Qy 1141 TTTATAAATACTAATTTGCCAGCAATTTGCTGCTAAAGAAAGGAGGTTATCTCTGAGAAA 1200  
Db 1132 -----GAGGAGGCTGAAGCTGAAGCTGAACCTTAAATTCATCAACACT 1173  
Qy 1201 AGAGAGGCTGAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGGAGGAGGAGGATGATCATG 1260  
Db 1174 ACTATGCTTCTATGCTGCTTAAGGAGGAGGTTTCTCTGAGAAAAGAGAGGCTGAG 1233  
Qy 1261 GCGCCCGGGGAGGAGGATAAATTTGGCTTCCCGGTGTCAACCGCCAGCAGTGCACGGAG 1320  
Db 1234 GCTTACGTAGAGTTGA----- 1249  
Qy 1321 AGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTTCCCGTGTGCTTCCACCCCATGGCC 1380  
Db 1250 ----- 1249  
Qy 1381 ATCAGAAACACTCAAGAGAGAGATGTCCTTTCTAACTAGTAGGCGTAGAATTTCCCTAGGG 1440

Db 1250 -----GCTCGGG 1256  
Qy 1441 CGGCCGCAATTAATTCGCTTAGACATGACTGTTTCTCAGTTCAAGTTGGCAGCTTACG 1500  
Db 1257 CGGCCGCAATTAATTCGCTTAGACATGACTGTTTCTCAGTTCAAGTTGGCAGCTTACG 1316  
Qy 1501 AGAAGACCGGCTCTGCTAGATTTCTAATCAAGAGGATGTCAAGATGCCATTTTCCCTGAGAG 1560  
Db 1317 AGAAGACCGGCTCTGCTAGATTTCTAATCAAGAGGATGTCAAGATGCCATTTTCCCTGAGAG 1376  
Qy 1561 ATGCAAGGCTTCAATTTTGTATGACTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1620  
Db 1377 ATGCAAGGCTTCAATTTTGTATGACTTTTATTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1436  
Qy 1621 GTCATTTTCTTCTCTGCTAGAGCTTCTCTCATGACGCTATCTCGCAGCTCATGAA 1680  
Db 1437 GTCATTTTCTTCTCTGCTAGAGCTTCTCTCATGACGCTATCTCGCAGCTCATGAA 1496  
Qy 1681 TATCTTGTGCTAGGGTTTGGGAAATCATTTCGAGTTGATGTTTTTCTTGTGATTTTCC 1740  
Db 1497 TATCTTGTGCTAGGGTTTGGGAAATCATTTCGAGTTGATGTTTTTCTTGTGATTTTCC 1556  
Qy 1741 ACTCCTCTTTCAGAGTACAGAAATTAAGTGAGAGTTCGTTTGTGCAAGCTTATCGATAA 1800  
Db 1557 ACTCCTCTTTCAGAGTACAGAAATTAAGTGAGAGTTCGTTTGTGCAAGCT-----A 1608  
Qy 1801 GCTTTAATGCGGTAGTTTATCAAGTTAAATTTGCTAAACGAGTCAGGCAACCGTGTATGAA 1860  
Db 1609 GCTTTAATGCGGTAGTTTATCAAGTTAAATTTGCTAAACGAGTCAGGCAACCGTGTATGAA 1668  
Qy 1861 AUCTAAATGCGCTCATGCTCGGACCGTCACCTCGGATGCTGTAGGANTAGG 1920  
Db 1669 AUCTAAATGCGCTCATGCTCGGACCGTCACCTCGGATGCTGTAGGANTAGG 1728  
Qy 1921 CTTGGTTATGCGGTACTGCGGGCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1980  
Db 1729 CTTGGTTATGCGGTACTGCGGGCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1788  
Qy 1981 CAGTCACTATGCGCTCTCTAGCGCTATATGCTGTTGATGCAATTTCTATGCGCACCCGT 2040  
Db 1789 CAGTCACTATGCGCTCTCTAGCGCTATATGCTGTTGATGCAATTTCTATGCGCACCCGT 1848  
Qy 2041 TCTCGGAGCACTGTCCGACCGCTTTGGCGCGCCGCGCAGTCTGCTCGCTTCTGCTACTTGG 2100  
Db 1849 TCTCGGAGCACTGTCCGACCGCTTTGGCGCGCCGCGCAGTCTGCTCGCTTCTGCTACTTGG 1908  
Qy 2101 AGCCACTATCGACTACGCGATCATGCGGACCAACACCCGCTCTGTGGATCTATCGAATCTA 2160  
Db 1909 AGCCACTATCGACTACGCGATCATGCGGACCAACACCCGCTCTGTGGATCTATCGAATCTA 1968  
Qy 2161 AATGTAAGTTAAATCTTAATAATTAAGTCCGAGTTTCTCCATACGAACTTAA 2220  
Db 1969 AATGTAAGTTAAATCTTAATAATTAAGTCCGAGTTTCTCCATACGAACTTAA 2028  
Qy 2221 CAGCATTTCGCTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2280  
Db 2029 CAGCATTTCGCTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2088  
Qy 2281 TATGTTTCACTCCCTCAGGAGTTACGTTTGTGAGTGAATTAATCTTCTGGAAGGTTGCGAG 2340  
Db 2089 TATGTTTCACTCCCTCAGGAGTTACGTTTGTGAGTGAATTAATCTTCTGGAAGGTTGCGAG 2148  
Qy 2341 TGTTRAATCCGCTGATTTGACGGGATATCCGTAAGTTCGCAAGTGTGGTTGCTGTAACCGG 2400  
Db 2149 TGTTRAATCCGCTGATTTGACGGGATATCCGTAAGTTCGCAAGTGTGGTTGCTGTAACCGG 2208  
Qy 2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGCAACAAACAAACAGATCCAGCGTGT 2460  
Db 2209 AGGAGTAATCTCCAACTCTCTGAGAGTAGGCAACAAACAAACAGATCCAGCGTGT 2268  
Qy 2461 GTACTTGCATCAATAAGAAAGCAATTTCTCGATTTTCGAGGATCAAGTGTTCAGGAGCGT 2520  
Db 2269 GTACTTGCATCAATAAGAAAGCAATTTCTCGATTTTCGAGGATCAAGTGTTCAGGAGCGT 2328

Qy	2521	ACTGATTGGACATTTCCAAAGCCCTGCTCGTAGGTTGCAACCGATAGGGTTGTAGAGTGTG	2580
Db	2329	ACTGATTGGACATTTCCAAAGCCCTGCTCGTAGGTTGCAACCGATAGGGTTGTAGAGTGTG	2388
Qy	2581	CAATACACTTTGGCGTACAAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTCTGTGAACAGCAT	2640
Db	2389	CAATACACTTTGGCGTACAAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTCTGTGAACAGCAT	2448
Qy	2641	CTTCAATTCCTGGCAAGCTCCTTGTCTCATATCGACAGCAACAGAAATCACCTGGGAAT	2700
Db	2449	CTTCAATTCCTGGCAAGCTCCTTGTCTCATATCGACAGCAACAGAAATCACCTGGGAAT	2508
Qy	2701	CAATACCAATGTTAGCTTTGAGACAGAGAGTCTCGAGCAACGAAATCTGGAATCAGCGTATT	2760
Db	2509	CAATACCAATGTTAGCTTTGAG - CAGAAGGTTCTGAGGCAACGAATCTGGAATCAGCGTATT	2567
Qy	2761	TATCAGCAATAACTAGAACTTTCAGAAGGCCACGACAGGCATGTCAAATACTACACAGGCGTG	2820
Db	2568	TATCAGCAATAACTAGAACTTTCAGAAGGCCACGACAGGCATGTCAAATACTACACAGGCGTG	2627
Qy	2821	ATGCTGTCATTTTGAACCATCATCTTGCAGCAGAGTAAAGAACTGGTTCCCTGGACCAATA	2880
Db	2628	ATGCTGTCATTTTGAACCATCATCTTGCAGCAGAGTAAAGAACTGGTTCCCTGGACCAATA	2687
Qy	2881	TTTTTGTACACTTTAGGAACAAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCCCTGGGCGC	2940
Db	2688	TTTTTGTACACTTTAGGAACAAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCCCTGGGCGC	2747
Qy	2941	CTCCTGCTAGGACGATACACTTAGCACCAACCTTGTGGGCAAACGTAGATGACTTCTGGGG	3000
Db	2748	CTCCTGCTAGGACGATACACTTAGCACCAACCTTGTGGGCAAACGTAGATGACTTCTGGGG	2807
Qy	3001	TAAGGGTACCATTCCCTTTAGTGGAGATGCAAAAACAATTTCTTTCGCAACCAAGCAACTT	3060
Db	2808	TAAGGGTACCATTCCCTTTAGTGGAGATGCAAAAACAATTTCTTTCGCAACCAAGCAACTT	2867
Qy	3061	TGGCAGGAAACCCAGCATCAGGGAAGTGGAGGAGCAATTGGCGTTCCACACAGGAATAT	3120
Db	2868	TGGCAGGAAACCCAGCATCAGGGAAGTGGAGGAGCAATTGGCGTTCCACACAGGAATAT	2927
Qy	3121	AGAGGCCAACTTTCTCAATAGTCTTTCGAAAAAGAGAGCAGACTACACACAGGCGCAAGTCT	3180
Db	2928	AGAGGCCAACTTTCTCAATAGTCTTTCGAAAAAGAGAGCAGACTACACACAGGCGCAAGTCT	2987
Qy	3181	CAACTTGCAGAGTCTCGGTTAGTTGAGCTTCATGGAAATTCCTCTGAAGTTATCTATAGAGA	3240
Db	2988	CAACTTGCAGAGTCTCGGTTAGTTGAGCTTCATGGAAATTCCTGAGCTTATCTATAGAGA	3047
Qy	3241	GATCAATGGCTCTCTTAAAGCTTATCTGGCAATTCGATAAGTTCTCTCGGGAAGAGGAGCTT	3300
Db	3048	GATCAATGGCTCTCTTAAAGCTTATCTGGCAATTCGATAAGTTCTCTCGGGAAGAGGAGCTT	3107
Qy	3301	CTAACACAGGTGCTTCAAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT	3360
Db	3108	CTAACACAGGTGCTTCAAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT	3167
Qy	3361	TGTCACCACTTTTGACGAACTATGTCGCAATTTGGTTTGACTTAATTCATTAATCTGTTCCG	3420
Db	3168	TGTCACCACTTTTGACGAACTATGTCGCAATTTGGTTTGACTTAATTCATTAATCTGTTCCG	3227
Qy	3421	TTTTTCTGGATAGGACGACGAAGGCACTCTTCAATTTCTTGTGAGGAGCGCTTAGAAAACGT	3480
Db	3228	TTTTTCTGGATAGGACGACGAAGGCACTCTTCAATTTCTTGTGAGGAGCGCTTAGAAAACGT	3287
Qy	3481	CAATTTTGCACAAATTCGAATCGACCTTCAGAGGAGCTTCTTTTAGGTTTTCGATCTCTCTT	3540
Db	3288	CAATTTTGCACAAATTCGAATCGACCTTCAGAGGAGCTTCTTTTAGGTTTTCGATCTCTCTT	3347
Qy	3541	TAGGTTGTTCTCTGGGTATACCTCGGCTTGGCATCTCTCTTTCTTCTAGTGACCTTTTAGGG	3600
Db	3348	TAGGTTGTTCTCTGGGTATACCTCGGCTTGGCATCTCTCTTTCTTCTAGTGACCTTTTAGGG	3407

Qy	3601	ACTTCAATATCCAGGTTTCTCTCCACCTCGTCCAAAGTCAACAGTACTTGGCACATCTAA	3660
Db	3408	ACTTCAATATCCAGGTTTCTCTCCACCTCGTCCAAAGTCAACAGTACTTGGCACATCTAA	3467
Qy	3661	CTAATGCCAAATAAATAAGTTCAGCACATTCACAGGCTATATCTTCCTTTGGATTTAGCTT	3720
Db	3468	CTAATGCCAAATAAATAAGTTCAGCACATTCACAGGCTATATCTTCCTTTGGATTTAGCTT	3527
Qy	3721	CTGCAAGTTCAATCAGCTTCTCTCCCTTAATTTTACGHTTC-AAACAAACTTCGTCTGTCATAT	3779
Db	3528	CTGCAAGTTCAATCAGCTTCTCTCCCTTAATTTTACGHTTC-AAACAAACTTCGTCTGTCATAT	3587
Qy	3780	AACCGTTTGTAATAAGACCTTCGTGAGCATTCCTTACAGATCCCAAGGTGGCTTCC	3839
Db	3588	AACCGTTTGTAATAAGACCTTCGTGAGCATTCCTTACAGATCCCAAGGTGGCTTCC	3646
Qy	3840	ATGGCTCTAAGACCTTTTCATTTGGCCAAACAGGAAGTCGTTTCAAAGTGAACAGAAACCA	3899
Db	3647	ATGGCTCTAAGACCTTTTCATTTGGCCAAACAGGAAGTCGTTTCAAAGTGAACAGAAACCA	3706
Qy	3900	ACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACATTCCAATTCGATACCC	3959
Db	3707	ACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACATTCCAATTCGATACCC	3766
Qy	3960	AGCAACTTTTGAGTGTCTCCAGATGTAGCACCTTTTATACCAAAACCGGTGACGACGAT	4019
Db	3767	AGCAACTTTTGAGTGTCTCCAGATGTAGCACCTTTTATACCAAAACCGGTGACGACGAT	3826
Qy	4020	TGTTAGACTCCAGTTTGTGTCTTATAGCCTTCGGAAATAGACTTTTGGACAGATACACC	4079
Db	3827	TGTTAGACTCCAGTTTGTGTCTTATAGCCTTCGGAAATAGACTTTTGGACAGATACACC	3886
Qy	4080	AGSCCACAACGAGTAATTTAGAAGTCAGCACCAAAAGTAGTCAATAGACCATTCGGGGCGG	4139
Db	3887	AGSCCACAACGAGTAATTTAGAAGTCAGCACCAAAAGTAGTCAATAGACCATTCGGGGCGG	3946
Qy	4140	TCAGTAGTCAAAAGCCCAACAAAATTTCACTGACAGGAACTTTTGTGACATCTTCAGAA	4199
Db	3947	TCAGTAGTCAAAAGCCCAACAAAATTTCACTGACAGGAACTTTTGTGACATCTTCAGAA	4006
Qy	4200	AGTTCGATATCAGTAGTCAATTCGCCAGCATCAATTAATGGGGATATACCAAGCAACA	4259
Db	4007	AGTTCGATATCAGTAGTCAATTCGCCAGCATCAATTAATGGGGATATACCAAGCAACA	4066
Qy	4260	GTGGAAGTCACATCTACCAACTTTGCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCG	4319
Db	4067	GTGGAAGTCACATCTACCAACTTTGCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCG	4126
Qy	4320	CAATTAGTGAACCTTTCAATTCGCCAGTCAGAGAAAGGACACAGCGATCTAGCA	4379
Db	4127	CCATTAGTGAACCTTTCAATTCGCCAGTCAGAGAAAGGACACAGCGATCTAGCA	4186
Qy	4380	TTAGCGGGCAGGATGCAACTTTTATCAACACAGGGTCTTATAGATAACCCCTAGCGCTGG	4439
Db	4187	TTAGCGGGCAGGATGCAACTTTTATCAACACAGGGTCTTATAGATAACCCCTAGCGCTGG	4246
Qy	4440	ATCATCCTTTGGACAACTCTTTCTGCAAAATCTAGTCCAAATCACTTCATGTATACCA	4499
Db	4247	ATCATCCTTTGGACAACTCTTTCTGCAAAATCTAGTCCAAATCACTTCATGTATACCA	4306
Qy	4500	TTATTTCTACAACTTGAGCAGTTGTGATCAGCTCCTCAAAATTTGTCCTCTGTAACGGAT	4559
Db	4307	TTATTTCTACAACTTGAGCAGTTGTGATCAGCTCCTCAAAATTTGTCCTCTGTAACGGAT	4366
Qy	4560	GACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTTAGTGAACCTTGATCAGTTGTGC	4619
Db	4367	GACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTTAGTGAACCTTGATCAGTTGTGC	4426
Qy	4620	AGCTGTGACGACATAGGGAACACCGCTTTTCTTACCAAACTCAAGGAATTTCAAC	4679
Db	4427	AGCTGTGACGACATAGGGAACACCGCTTTTCTTACCAAACTCAAGGAATTTCAAC	4486
Qy	4680	TCGCAACCACTTTCGCTATCGAGTAGCAAGGAAATGTCTATCTTGAAGTCGACAGTGA	4739

Db	4487	TCTGCAACACTTGGTATGACGATGAGTAAAGTCAATCTTGAAGTCGACAGTGA	4546	5567	ATATGAATAAATTGCAGTTTTCATTTGATGCTCGATGAGTTTTTCTTAATCAGAAATGGTTA	5626	
Qy	4740	GTGTAGTCTTGAGAAATTCGAGCCGTATTTTATTATCAGTCAAGTCAATCAGGAG	4799	4984	-----	4983	
Db	4547	GTGTAGTCTTGAGAAATTCGAGCCGTATTTTATTATCAGTCAATCAGGAG	4606	Db	5627	ATTGTTTAAACATGGCAGAGCAATTAACCTGACTTGAAGTGGGACCGCGGCTTTGTGTAAT	5686
Qy	4800	ATCCTCTACGCCGACGACATCGTGGCGG	4827	4984	-----	4983	
Db	4607	ATCCTCTACGCCGACGACATCGTGGCGGATGGTGTCTGGGTGAACAGGTTTATTCCTGGC	4666	Db	5687	AAATCGAACTTTTGCTGAGTTGAAGATCAGATCAGCATCTTCCCGACAAACGACAGCG	5746
Qy	4828	-----	4827	4984	-----	4983	
Db	4667	ATCCACTAAATATAGCATGCGAAAGCCACGTTGTCTCAAAATCTCTGATGTACATTG	4726	Db	5747	TTCCGTGCGAAAGCAAAAGTTTCAAAATCACCAACTGGTCCACTACAAACAAAGCTCTCAT	5806
Qy	4828	-----	4827	4984	-----	4983	
Db	4727	CACAAGATAAAATATATCATCATGAACAAATAAAATCTGTCTGTACATAACAGTAATA	4786	Db	5807	CAACCGTGGCTCCCTCCTCCTTCTGGCTGATGATGGGGGATTCAGGCCCTGGTATGAGTC	5866
Qy	4855	CAGTGGCGTTGTGGC	4872	4984	-----	5026	
Db	4787	CAAGGGGTGTATGAGCCATATTCACCGGAAACGTCTTTGTCGAGGCCGGATTAAT	4846	Qy	5867	AGCAACACCTTCTTACGAGGAGAGCTCAGCGCGCGGATTCCTTGGCGCGCGGTCT	5926
Qy	4873	CCTATATCGCGACATCACCGATGGGAAGATCGGGCTCGC	4913	Db	5027	CAACGGCTCAACCTACTACTCGGCTGCTTCTAATGACAGAGTCGCATAAAGGAGAGCG	5086
Db	4847	CCAACTGATGCTGATTTATATGGGTATATGGCTCGCGATAAATGTCGGCAATCAG	4906	Qy	5927	-AACGGCTCAACCTAC - GCGTGGCTGCTTCTTAATGACAGAGTCGCATAAAGGAGAGCG	5984
Qy	4914	-----	4913	Db	5087	TCAGTATCTATGATTGGAAGTATGGGAATGGTATACCCGCAATCTTTCAGTGTCTTGAG	5146
Db	4907	GTGCAACAATCTATCGATTGTATGGAAGCCGATGCGCCAGAGTTGTTCTGAACATG	4966	Qy	5985	TCAGTATCTATGATTGGAAGTATGGGAATGGTATACCCGCAATCTTTCAGTGTCTTGAG	6044
Qy	4914	-----	4913	Db	5147	GTCTCCTATCAGATTATGCCCACTAAAGCAACCGGAGGAGAGATTTTATGGTAAATTT	6104
Db	4967	GCAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAAGTAACTGGCTGACGG	5026	Qy	5207	CTCTGACTTTTGTCTCAGTAGACTCGAACTGAGACTATCTCGGTTATGACAGCAGA	5266
Qy	4914	-----	4913	Db	6105	CTCTGACTTTTGTCTCAGTAGACTCGAACTGAGACTATCTCGGTTATGACAGCAGA	6164
Db	5027	AATTTATGCTCTTCCGACCATCAAGCATTTTATCCGTACTCTCTGATGATGATGTTAC	5086	Qy	5267	AATGTCTTCTTGAGACAGTAAATGAAGTCCACCAATAAAGAAATCTTGTATACAGG	6224
Qy	4914	-----	4913	Db	5327	ACAAACTCTTCTTTCGAACTTTTTCGGTGCCTTGAACATATAAAATGTAGAGTGGATAT	5386
Db	5087	TCACCACTCGCATCCAGGGAACACAGCATTCACAGTATTAGAGAATATCTGATTCAG	5146	Qy	6225	ACAAACTCTTCTTTCGAACTTTTTCGGTGCCTTGAACATATAAAATGTAGAGTGGATAT	6284
Qy	4914	-----	4913	Db	5387	GTGGGTAGGAATGGAGCGGCAAAATGCTTACCTTTCGACCTTCAAGAGGTATGAGG	5446
Db	5147	GTCAAAATATTCTGATCGCTGCGAGTGTCTTCCGCGGTTGATTCGATTCCTGTTT	5206	Qy	6285	GTGGGTAGGAATGGAGCGGCAAAATGCTTACCTTTCGACCTTCAAGAGGTATGAGG	6344
Qy	4914	-----	4913	Db	5447	TTGTAGATAGTATGACCACTTTCAGTGACAACTTCTGATTTTTCGTTCAAACTCCGA	5506
Db	5207	GTAAATGTCTTTTAAACAGCATCGGTATTTTCGTCTCGCTCAGCGCAATCAGAAATGA	5266	Qy	6345	TTGTAGATAGTATGACCACTTTCAGTGACAACTTTCGATTTTTCGTTCAAACTCCGA	6404
Qy	4914	-----	4913	Db	5507	ATCCAGAGAAATCAAACTTGTGTCTTACTTATGATCCAAAGCCAGTCCGCTTGAAC	5566
Db	5267	ATAACGTTTGTGATGCGAGTATTTGATGACAGCGTAAATGGCTGGCTGTTGAAC	5326	Qy	6405	ATCCAGAGAAATCAAACTTGTGTCTTACTTATGATCCAAAGCCAGTCCGCTTGAAC	6464
Qy	4914	-----	4913	Db	5567	GACAAATAGTGTGCTGTTTTCAGTGTATGATTAATCTAGTCTTGTATCT	5626
Db	5327	AACTCTGGAAGAAATGCATAAACTTTTGCATTTCTACCGGATTCAGTGTCTCATG	5386	Qy	6465	GACAAATAGTGTGCTGTTTTCAGTGTATGATTAATCTAGTCTTGTATCT	6524
Qy	4914	-----	4913	Db	5627	AAATAATCTTTCAGCAGCAAGCGGATAAATACCCAAATCTAAAACTCTTTTAAACCGTTA	5686
Db	5387	GTGATTTCTCATTGATTAACCTTATTTTGAAGGGGAAATTAATGAGTGTGATG	5446	Qy	6525	AAATAATCTTTCAGCAGCAAGCGGATAAATACCCAAATCTAAAACTCTTTTAAACCGTTA	6584
Qy	4972	GGGAGCTGTTGG	4983	Db	5687	AAAGGACAAGTATGCTGCTCTGTTTAAACCCCAATCAGTCTGATCTGATCTCATC	5746
Db	5447	TTGACGAGTGGGAATCGCAGACCGATACAGGATCTTGCCTCTTATGGAATGCTCG	5506	Qy	6585	AAAGGACAAGTATGCTGCTCTGTTTAAACCCCAATCAGTCTGATCTGATCTCATC	6644
Qy	4984	-----	4983	Db	5747	AACCTGAGGCGCATATCTTGTGTTTGTAGGAAATTTGCGAGATCGGATATCGAGAAAG	5806
Db	5507	GTGAGTTTCTCCTTCAATACAGAAACGGCTTTTCAAAATATGATGATTAATCTCTG	5566	Qy	6645	AACCTGAGGCGCATATCTTGTGTTTGTAGGAAATTTGCGAGATCGGATATCGAGAAAG	6704
Qy	4984	-----	4983	Db			

QY 5807 GTACGCTGATTTTAAACGTGAATTTATCTCAAGATCTCTGCTCGCGCTTTCCGGTGAT 5866  
Db 6705 GTACGCTGATTTTAAACGTGAATTTATCTCAAGATCTCTGCTCGCGCTTTCCGGTGAT 6764  
QY 5867 GACGGTGAACCTCTGACACATGACGCTCCCGGAGAGCGGTACAGAGTCTGTCTGAAGCG 5926  
Db 6765 GACGGTGAACCTCTGACACATGACGCTCCCGGAGAGCGGTACAGAGTCTGTCTGAAGCG 6824  
QY 5927 GATGCCGGGAGACACAGCCCGGTACAGGCCCGCTCAGCGGGTGTGGCGGGTGTCCGGGC 5986  
Db 6825 GATGCCGGGAGACACAGCCCGGTACAGGCCCGCTCAGCGGGTGTGGCGGGTGTCCGGGC 6884  
QY 5987 GCAGCCATGACCCAGTACAGTAGCGAGTGTATCTGGCTTAACTATCCGGCAT 6046  
Db 6885 GCAGCCATGACCCAGTACAGTAGCGAGTGTATCTGGCTTAACTATCCGGCAT 6944  
QY 6047 CAGACGATTTGACTGAGAGTGACACATATGCGGTGTGAATAACCGACAGATGCGTAA 6106  
Db 6945 CAGACGATTTGACTGAGAGTGACACATATGCGGTGTGAATAACCGACAGATGCGTAA 7004  
QY 6107 GGAGAAATACCGCATCAGGCGCTCTCCGCTTCTCGCTCACGTCTCGTGGCTCGG 6166  
Db 7005 GGAGAAATACCGCATCAGGCGCTCTCCGCTTCTCGCTCACGTCTCGTGGCTCGG 7064  
QY 6167 TCGTTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAGGGCGTAAATACGGTTATCCACAG 6226  
Db 7065 TCGTTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAGGGCGTAAATACGGTTATCCACAG 7124  
QY 6227 AATCAGGGGTAACCGCAGGAAGACATGTGACGAAAGCCAGCAAGCCAGCAAGC 6286  
Db 7125 AATCAGGGGTAACCGCAGGAAGACATGTGACGAAAGCCAGCAAGCCAGCAAGC 7184  
QY 6287 GTAAAAGCGCGCTGTGCTGGCTTTTCCATAGGCTCCGCCCCCTTGACGAGCATCACA 6346  
Db 7185 GTAAAAGCGCGCTGTGCTGGCTTTTCCATAGGCTCCGCCCCCTTGACGAGCATCACA 7244  
QY 6347 AATATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGAGCTATAAGATACAGCGGT 6406  
Db 7245 AATATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGAGCTATAAGATACAGCGGT 7304  
QY 6407 TTCCCTCGGAAGCTCCCTCTGCGCTCTCTGCTTCGACCCCTGCGCTTACCGGATACC 6466  
Db 7305 TTCCCTCGGAAGCTCCCTCTGCGCTCTCTGCTTCGACCCCTGCGCTTACCGGATACC 7364  
QY 6467 TGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATC 6526  
Db 7365 TGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATC 7424  
QY 6527 TCAGTTCGGTGTAGGTCTGCTCCAGCTGGGCTGTGTCAGCAAGCCCGCTTACG 6586  
Db 7425 TCAGTTCGGTGTAGGTCTGCTCCAGCTGGGCTGTGTCAGCAAGCCCGCTTACG 7484  
QY 6587 CCGACCGCTGCGCTTATCCGGTAACATCGTCTTGAGTCCAAACCCGGTAAGACAGCT 6646  
Db 7485 CCGACCGCTGCGCTTATCCGGTAACATCGTCTTGAGTCCAAACCCGGTAAGACAGCT 7544  
QY 6647 TATGCCCATGCGACGACCACTGGTAACAGGATTACAGAGCGAGGTATGTAGGCGGTG 6706  
Db 7545 TATGCCCATGCGACGACCACTGGTAACAGGATTACAGAGCGAGGTATGTAGGCGGTG 7604  
QY 6707 CTACAGAGTCTTGAAGTGTGGCTTAACCTACGCTTACCTAGAGGACAGTATTTGGTA 6766  
Db 7605 CTACAGAGTCTTGAAGTGTGGCTTAACCTACGCTTACCTAGAGGACAGTATTTGGTA 7664  
QY 6767 TCTCGCTCTGCTGAAGCGAGTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCA 6826  
Db 7665 TCTCGCTCTGCTGAAGCGAGTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCA 7724  
QY 6827 AACAAACACCGCTGGTAGCGGTGGTGTGTTTGTGCAAGCAGCAGATACCGCGAGAA 6886  
Db 7725 AACAAACACCGCTGGTAGCGGTGGTGTGTTTGTGCAAGCAGCAGATACCGCGAGAA 7784

QY 6887 AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACG 6946  
Db 7785 AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACG 7844  
QY 6947 AAAACTCAGTTAAGGGATTTTGGTCAAGAGATTAACAAAAAGGATCTTACCTAGATCC 7006  
Db 7845 AAAACTCAGTTAAGGGATTTTGGTCAAGAGATTAACAAAAAGGATCTTACCTAGATCC 7904  
QY 7007 TTTTAAATTTAAATTTAAATTTAAATCAATCTAAAGTATATATAGTAATAACTTGGTCTG 7066  
Db 7905 TTTTAAATTTAAATTTAAATTTAAATCAATCTAAAGTATATATAGTAATAACTTGGTCTG 7964  
QY 7067 ACAGTTACCAATTCGCTTAATCAGTAGGAGCACCTATCTCAGCGATCTGTCTATTTGGTTCAT 7126  
Db 7965 ACAGTTACCAATTCGCTTAATCAGTAGGAGCACCTATCTCAGCGATCTGTCTATTTGGTTCAT 8024  
QY 7127 CCATAGTTGCTGACTCCCGTCTGTGTAGATAACTACGATACGGAGGGCTTACCATCTG 7186  
Db 8025 CCATAGTTGCTGACTCCCGTCTGTGTAGATAACTACGATACGGAGGGCTTACCATCTG 8084  
QY 7187 GCCCAGTGTCTCAATGATACCGCGAGACCCACGCTCACCGCTCCAGATTTATCAGCAA 7246  
Db 8085 GCCCAGTGTCTCAATGATACCGCGAGACCCACGCTCACCGCTCCAGATTTATCAGCAA 8144  
QY 7247 TAAACACGACGCGGAGGCGAGCGAGAGTGTCTGCAACTTTTATCCGCTTCCA 7306  
Db 8145 TAAACACGACGCGGAGGCGAGCGAGAGTGTCTGCAACTTTTATCCGCTTCCA 8204  
QY 7307 TCCAGTCTATTAATTTGCGGGAAGCTAGAGTAAGTGTTCGCCAGTAAATAGTTTCG 7366  
Db 8205 TCCAGTCTATTAATTTGCGGGAAGCTAGAGTAAGTGTTCGCCAGTAAATAGTTTCG 8264  
QY 7367 GCAAGCTTTGTCCTATCTCAGCATCTGTGGTGTACGCTCGCTTTCGTTATGGCTT 7426  
Db 8265 GCAAGCTTTGTCCTATCTCAGCATCTGTGGTGTACGCTCGCTTTCGTTATGGCTT 8324  
QY 7427 CATTCAGCTCCGTTCCCAACGATCAGGCGAGTACATGATCCCCCATGTTGTCRAAA 7486  
Db 8325 CATTCAGCTCCGTTCCCAACGATCAGGCGAGTACATGATCCCCCATGTTGTCRAAA 8384  
QY 7487 AAGCGGTAGTCTCTCGCTCTCCGATCTGTGTCAGAAAGTAAAGTTGGCCCGAGTGTAT 7546  
Db 8385 AAGCGGTAGTCTCTCGCTCTCCGATCTGTGTCAGAAAGTAAAGTTGGCCCGAGTGTAT 8444  
QY 7547 CACTCATGGTTATGCGACACATGCAATTTCTCTTACTGTCTATGCCATCCGTAAGTGT 7606  
Db 8445 CACTCATGGTTATGCGACACATGCAATTTCTTACTGTCTATGCCATCCGTAAGTGT 8504  
QY 7607 TTTCTGTGACTGGTGAAGTCAACCAAGTCAATCTGAGAATAGTGTATGGCGGACCGA 7666  
Db 8505 TTTCTGTGACTGGTGAAGTCAACCAAGTCAATCTGAGAATAGTGTATGGCGGACCGA 8564  
QY 7667 GTTGTCTTTCGCGCGCTCAACACGGGATAATACCGGCCACATAGCAGAACTTTAAAG 7726  
Db 8565 GTTGTCTTTCGCGCGCTCAACACGGGATAATACCGGCCACATAGCAGAACTTTAAAG 8624  
QY 7727 TGCTCATCATTTGGAAACGTTCTTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTGA 7786  
Db 8625 TGCTCATCATTTGGAAACGTTCTTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTGA 8684  
QY 7787 GATCCAGTTTCGATGTAAACCCACTCGTGACCCACTGATCTTCAGCATCTTTACTTTCA 7846  
Db 8685 GATCCAGTTTCGATGTAAACCCACTCGTGACCCACTGATCTTCAGCATCTTTACTTTCA 8744  
QY 7847 CCAGGTTTCTGGGTGACAAAACAGAGGCGAAATTCGCGAAAAGGATAGG 7906  
Db 8745 CCAGGTTTCTGGGTGACAAAACAGAGGCGAAATTCGCGAAAAGGATAGG 8804  
QY 7907 GCACACGAAATGTGAATACTACTCTTCTCTTTTCAATATTATTGAAGCATTTATC 7966  
Db 8805 GCACACGAAATGTGAATACTACTCTTCTCTTTTCAATATTATTGAAGCATTTATC 8864  
QY 7967 AGGTTATTTCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 8026

8865 AGGTTATTTCTCATAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAAATAG 8924  
8027 GGGTTCCGGGCACATTTCCCGGAAAAGTGCACCTGACGCTTAAGAAACCAATTATTATCA 8086  
8925 GGGTTCCGGGCACATTTCCCGGAAAAGTGCACCTGACGCTTAAGAAACCAATTATTATCA 8984  
8087 TGACATTAACCTATAAAAAATAGGGGTATACAGAGGCCCTTTCCTTCAAGAAATTAATTC 8146  
8985 TGACATTAACCTATAAAAAATAGGGGTATACAGAGGCCCTTTCCTTCAAGAAATTAATTC 9044  
8147 TCATGTTTGACAGCTTATCATCGATAAGCTGACATCATGTTGGTATTTGAAATAGACGCA 8206  
9045 TCAAGTTTGACAGCTTATCATCGATAAGCTGACATCATGTTGGTATTTGAAATAGACGCA 9104  
8207 GATCGGAACACTGAAAAATAACAGTTATTATTTCG 8241  
9105 GATCGGAACACTGAAAAATAACAGTTATTATTTCG 9139

RESULT 12  
AX146621 10462 bp DNA linear PAT 31-MAY-2001  
LOCUS Sequence 83 from Patent WO0134654.  
DEFINITION AX146621  
ACCESSION AX146621  
VERSION AX146621.1 GI:14285014  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 Strauch, K.  
Hedgehog fusion proteins and uses  
Patent: WO 0134654-A 83 17-MAY-2001;  
BIOGEN, INC. (US)  
LOCATION/Qualifiers  
1..10462  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 69.3%; Score 5709.8; DB 6; Length 10462;  
Best Local Similarity 77.4%; Pred. No. 0;  
Matches 8096; Conservative 0; Mismatches 149; Indels 2221; Gaps 8;

QY 2 GATTAACATCAAGACGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 61  
DB 1 GATTAACATCAAGACGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 60  
QY 62 TCATTCTCACATAAGTGCACCAACGCAAGAGGGGATACACTAGCAGCAGACCGTT 121  
DB 61 TCATTCTCACATAAGTGCACCAACGCAAGAGGGGATACACTAGCAGCAGACCGTT 120  
QY 122 GCAAGCGAGGACCTCCACTCTCTCTCAACACCGACTTTTGGCCATCGAAAACCA 181  
DB 121 GCAAGCGAGGACCTCCACTCTCTCTCAACACCGACTTTTGGCCATCGAAAACCA 180  
QY 182 GCCAGTTATGGGCTTGATGGAGCTGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 241  
DB 181 GCCAGTTATGGGCTTGATGGAGCTGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 240  
QY 242 CACCATGACTTTATTAGCCTGCTACTCTGGCCCCCTGGGAGGTTTCAATGTTTAT 301  
DB 241 CACCATGACTTTATTAGCCTGCTACTCTGGCCCCCTGGGAGGTTTCAATGTTTAT 300  
QY 302 TTCGGAATGCAACAGCTCGCATTTACACCGCAATCATCTCCAGATGGGCTTTCTGA 361  
DB 301 TTCGGAATGCAACAGCTCGCATTTACACCGCAATCATCTCCAGATGGGCTTTCTGA 360  
QY 362 GTGTGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAAACTGACAGTTTAAACGCTG 421

DB 361 GTGTGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAAACTGACAGTTTAAACGCTG 420  
QY 422 TCTTCGAACCTTAATATGACAAAGCGTGTATCTCTCCCAAGATGAACTAAGTTTGGTTCGT 481  
DB 421 TCTTCGAACCTTAATATGACAAAGCGTGTATCTCTCCCAAGATGAACTAAGTTTGGTTCGT 480  
QY 482 TGAATGTCTAACCGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCCATACCGTTTGTG 541  
DB 481 TGAATGTCTAACCGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCCATACCGTTTGTG 540  
QY 542 TTGTTTGGTATTGATGAGCAATGCTCAAAATATATCTCATTAATGCTTAGCGGAGTCTC 601  
DB 541 TTGTTTGGTATTGATGAGCAATGCTCAAAATATATCTCATTAATGCTTAGCGGAGTCTC 600  
QY 602 TCTATCGCTTCTGAACCGCGTGCACCTGTGCCGAAAACGCAAAATGGGAAAACACCCGCTT 661  
DB 601 TCTATCGCTTCTGAACCGCGTGCACCTGTGCCGAAAACGCAAAATGGGAAAACACCCGCTT 660  
QY 662 TTTGGATGATATGATGCTCTCCACATTTGATGTTTCCAGATTTCTGGTGGGAATACTG 721  
DB 661 TTTGGATGATATGATGCTCTCCACATTTGATGTTTCCAGATTTCTGGTGGGAATACTG 720  
QY 722 CTGATAGCCTTAACGTTTCATGATCAAAATTTAACTGTTCTTAACCCCTACTTTGACAGCAATA 781  
DB 721 CTGATAGCCTTAACGTTTCATGATCAAAATTTAACTGTTCTTAACCCCTACTTTGACAGCAATA 780  
QY 782 TATAACAGAGGAAGCTGCCCTGTCTTAAACCTTTTATTTTATCATCATTTATTAGCTTTA 841  
DB 781 TATAACAGAGGAAGCTGCCCTGTCTTAAACCTTTTATTTTATCATCATTTATTAGCTTTA 840  
QY 842 CTTTCATAATTTGGAGCTGTTCCAAATTTGACAGCTTTTGAATTTTAACGACTTTTAAACGAC 901  
DB 841 CTTTCATAATTTGGAGCTGTTCCAAATTTGACAGCTTTTGAATTTTAAACGACTTTTAAACGAC 900  
QY 902 AACTTCAGAGATCAAAAAACAACTAAATTTATTCGAGAGATCCAAACGATGAGATTTCCCTT 961  
DB 901 AACTTCAGAGATCAAAAAACAACTAAATTTATTCGAGAGATCCAAACGATGAGATTTCCCTT 960  
QY 962 CAATTTTACTGCACTTTTATTCGAGCATCTCCGCAATAGTCTGCCAGTCAACACTA 1021  
DB 961 CAATTTTACTGCACTTTTATTCGAGCATCTCCGCAATAGTCTGCCAGTCAACACTA 1020  
QY 1022 CAACAGATGAAAGCGCAAAATTCGCGCTGAAGCTGTCATCGGTTTACTCAGATTTAG 1081  
DB 1021 CAACAGATGAAAGCGCAAAATTCGCGCTGAAGCTGTCATCGGTTTACTCAGATTTAG 1080  
QY 1082 AAGGGATTTTCGATGTTGCTGTTTCCCAATTTTCCAAAGCACAATAACGGGTTATTGT 1141  
DB 1081 AAGGGATTTTCGATGTTGCTGTTTCCCAATTTTCCAAAGCACAATAACGGGTTATTGT 1140  
QY 1142 TTATAATACTACTATTGCCAGCATTTGCTGCTAAAGAGAGGGGTATCTCTCGAGAAA 1201  
DB 1141 TTATAATACTACTATTGCCAGCATTTGCTGCTAAAGAGAGGGGTATCTCTCGAGAAA 1200  
QY 1202 GAGAGG----- 1207  
DB 1201 GATCGGACCGGCGAGGGGTTTCGGGAAGAGGAGGCACCCCAAAAGCTGACCCCTTTAG 1260  
QY 1208 ----- 1207  
DB 1261 CCTACAGCAGTTTATCCCAATTTGCGCGAGAGAACCCCTAGCGCCAGCGGAAGGTATG 1320  
QY 1208 ----- 1207  
DB 1321 AAGGAAAGATCTCCAGAACTCCGAGGGATTTAAGNACTCACCCCAATTACAAACCCCG 1380  
QY 1208 ----- 1207  
DB 1381 ACATCATATTTAAGGATGAAGAAAAACACCGGAGGAGAGGCTGATGACTCAGAGGTGTA 1440  
QY 1208 ----- 1207  
DB 1441 AGGACAAAGTTGAACGCTTTGGCCATCTCGGTGATGAACAGTGGCCAGGAGTGAACACTGC 1500





3661	Db	CAGAATCACTCGGAATCAATACCATGTTTCAGCTTGAG-CAGAAAGTCTGAGCCAAACGAA	3719
2744	Qy	ATCTCGATCAGCGTATTTATTCAGCATAACTAGAACTTCAGAAAGCCAGCAGGATGTC	2803
3720	Db	ATCTCGATCAGCGTATTTATTCAGCAATACTAGAACTTCAGAAAGCCAGCAGGATGTC	3779
2804	Qy	AATACTACACAGCGCTGATGTGTCAATTTTGAACCATCATCTTTGGCAGCAGTAAACGAATG	2863
3780	Db	AATACTACACAGCGCTGATGTGTCAATTTTGAACCATCATCTTTGGCAGCAGTAAACGAATG	3839
2864	Qy	GTITTCCTGGACCAAAATATTTTGTACACATTAGAAACAGTTTCTGTGTCCTGTAGCCATAGC	2923
3840	Db	GTITTCCTGGACCAAAATATTTTGTACACATTAGAAACAGTTTCTGTGTCCTGTAGCCATAGC	3899
2924	Qy	AGCTACTGCTGGCGCGCTCCTCTAGCAGATACATCTTAGCACCAACCTTGTGGGCAAC	2983
3900	Db	AGCTACTGCTGGCGCGCTCCTCTAGCAGATACATCTTAGCACCAACCTTGTGGGCAAC	3959
2984	Qy	GTAGATGACCTTCGGGGTAAGGTAACAATCTTTAGTGGAGATGCAAAACAAATTTTC	3043
3960	Db	GTAGATGACCTTCGGGGTAAGGTAACAATCTTTAGTGGAGATGCAAAACAAATTTTC	4019
3044	Qy	TTTGGCAACAGCAACTTTGGCAGAAACACCCAGCATCAGGAAAGTGGAAAGCAGAAATTCG	3103
4020	Db	TTTGGCAACAGCAACTTTGGCAGAAACACCCAGCATCAGGAAAGTGGAAAGCAGAAATTCG	4079
3104	Qy	GGTTCACACAGGAATATAGAGGCCAATTTCTCAATAGGTCCTTGCGAAAACGAGAGCAGAC	3163
4080	Db	GGTTCACACAGGAATATAGAGGCCAATTTCTCAATAGGTCCTTGCGAAAACGAGAGCAGAC	4139
3164	Qy	TACACAGGGCAAGTCTCAACTTTGCAACGCTCCGTTAGTTAGCTTCATGGAAATTTCTT	3223
4140	Db	TACACAGGGCAAGTCTCAACTTTGCAACGCTCCGTTAGTTAGCTTCATGGAAATTTCTT	4199
3224	Qy	GAGCTTATCTATAGAGAGATCAATGGCTCTCTTAAACGTTATCTGGCAATTCGATATAGTTC	3283
4200	Db	GAGCTTATCTATAGAGAGATCAATGGCTCTCTTAAACGTTATCTGGCAATTCGATATAGTTC	4259
3284	Qy	CTCTGGAAAGGAGCTTCTAACACAGGTGCTCTCAAGCGACCTCCATCAAACTTGGCAGT	3343
4260	Db	CTCTGGAAAGGAGCTTCTAACACAGGTGCTCTCAAGCGACCTCCATCAAACTTGGCAGT	4319
3344	Qy	TAGTTCTTAAAGGGCTTTGTCAACATTTTGACCAACAATTCGACAAATTCGTTTGACTAA	3403
4320	Db	TAGTTCTTAAAGGGCTTTGTCAACATTTTGACCAACAATTCGACAAATTCGTTTGACTAA	4379
3404	Qy	TTCCATTAATCTGTTCCGTTTCTGGATAGACAGCAGAGGGCATCTCAATTTCTTTGTA	3463
4380	Db	TTCCATTAATCTGTTCCGTTTCTGGATAGACAGCAGAGGGCATCTCAATTTCTTTGTA	4439
3464	Qy	GGAGGCTTAGAAAACGTCAAATTTTGGCAAAATCAATACAGCCTTCAGAGGGGACTTCTTT	3523
4440	Db	GGAGGCTTAGAAAACGTCAAATTTTGGCAAAATCAATACAGCCTTCAGAGGGGACTTCTTT	4499
3524	Qy	AGGTTTGGATTCCTCTTTAGGTTGTTCTGTTGTTATCTCTGGCTTGGCATCTCCTTTCCCT	3583
4500	Db	AGGTTTGGATTCCTCTTTAGGTTGTTCTGTTGTTATCTCTGGCTTGGCATCTCCTTTCCCT	4559
3584	Qy	TCTAGTGACCTTTAGGACTTCATATCCAGGTTTCTCTCCACCTCGTCCAAACGTCACACC	3643
4560	Db	TCTAGTGACCTTTAGGACTTCATATCCAGGTTTCTCTCCACCTCGTCCAAACGTCACACC	4619
3644	Qy	GTACTTGGCACAATCACTAATAAGCAAAATAAAATAGTACGACATTTCCCGAGGCTATATC	3703
4620	Db	GTACTTGGCACAATCACTAATAAGCAAAATAAAATAGTACGACATTTCCCGAGGCTATATC	4679
3704	Qy	TTCCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTAAATTTTAGCGTTC-AA	3762
4680	Db	TTCCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTAAATTTTAGCGTTC-AA	4739
3763	Qy	AAACTTCCTCGTCAAAATAACCGTTTGGTATAGAAACCTTCTGGAGCAATTCCTTACGAT	3822

4740	Db	AAACTTCGTCGTCAAATAAACCCTTTGGTATAGAAACCTTCTCGAGCATTCGCTCTTACGAT	47799
3823	Qy	CCCAACAGGTGGCTTCCATGGCTCTAAGACCCCTTTGATTTGGCCAAACACAGGAAGTGGCTT	3882
4800	Db	CCCAACAGGT -GCTTCCATGGCTCTAAGACCCCTTTGATTTGGCCAAACACAGGAAGTGGCTT	4858
3883	Qy	CCAAGTGCAGAAACAAACAAACCTGTTGTTTCAACACCAAAATTTCAAGCAGCTCTCCATCAC	3942
4859	Db	CCAAGTGCAGAAACAAACAAACCTGTTGTTTCAACACCAAAATTTCAAGCAGCTCTCCATCAC	4918
3943	Qy	AAATCCAATTCGATATACCCAGCAACTTTTGGATTTGCTCCAGATGTAGCACCTTTTATACCA	4002
4919	Db	AATCCAATTCGATATACCCAGCAACTTTTGGATTTGCTCCAGATGTAGCACCTTTTATACCA	4978
4003	Qy	AACCGTGCAGACAGATTTGTTAGACTCCAGTTTGTGTCCTTATAGCCTCCGGAAATAGACT	4062
4979	Db	AACCGTGCAGACAGATTTGTTAGACTCCAGTTTGTGTCCTTATAGCCTCCGGAAATAGACT	5038
4063	Qy	TTTTGGACGAGTACACACAGGCCAAACGAGTAATTAGAAGAGTCAGCCACCAAAAGTAGTGA	4122
5039	Db	TTTTGGACGAGTACACACAGGCCAAACGAGTAATTAGAAGAGTCAGCCACCAAAAGTAGTGA	5098
4123	Qy	ATAGACCATCGGGCGGTGAGTAGTCAAGACGCGCAACAAATTTCACTGACAGGGAACT	4182
5099	Db	ATAGACCATCGGGCGGTGAGTAGTCAAGACGCGCAACAAATTTCACTGACAGGGAACT	5158
4183	Qy	TTTTTGACATCTTCAGAAAGTTCGTATTCACTAGTCAATTCGCCAGCATCAATTAATGGGGA	4242
5159	Db	TTTTTGACATCTTCAGAAAGTTCGTATTCACTAGTCAATTCGCCAGCATCAATTAATGGGGA	5218
4243	Qy	TTATACAGAAACAGTGGAGTCACTCTACCAACTTTTGGGTCTCAGAAAAAGCAT	4302
5219	Db	TTATACAGAAACAGTGGAGTCACTCTACCAACTTTTGGGTCTCAGAAAAAGCAT	5278
4303	Qy	AAACAGTCTTACTACCGCCATTAGTGAACCTTTTCAATCGCCAGTCGAGAGAAAAAG	4362
5279	Db	AAACAGTCTTACTACCGCCATTAGTGAACCTTTTCAATCGCCAGTCGAGAGAAAAAG	5338
4363	Qy	GCACAGCATATAGCATTTAGCGGCAAGGATGCAACTTTATCAACAGGGTCTCTATAGA	4422
5339	Db	GCACAGCATATAGCATTTAGCGGCAAGGATGCAACTTTATCAACAGGGTCTCTATAGA	5398
4423	Qy	TAAACCTTAGCGCTGGGATCACTCTTTGGCACTCTTTCTGCCAAATCTAGGTCCTAAA	4482
5399	Db	TAAACCTTAGCGCTGGGATCACTCTTTGGCACTCTTTCTGCCAAATCTAGGTCCTAAA	5458
4483	Qy	TCACCTCATTTGATACCATTTATGTACAACTTGACAGTGTGCGATCAGTCTCTCAAAAT	4542
5459	Db	TCACCTCATTTGATACCATTTATGTACAACTTGAGCAAGTGTGCGATCAGTCTCTCAAAAT	5518
4543	Qy	GGTCTCTGTAACGGATGACTCAACTTTGCACATTTAACTTGAAGCTCAGTCGATTTAGTGA	4602
5519	Db	GGTCTCTGTAACGGATGACTCAACTTTGCACATTTAACTTGAAGCTCAGTCGATTTAGTGA	5578
4603	Qy	ACTTTGATCAGGTTGTGACGTTGGTTCAGCAGCATAGGGAACACAGGCTTTTCTTCAAAAC	4662
5579	Db	ACTTTGATCAGGTTGTGACGTTGGTTCAGCAGCATAGGGAACACAGGCTTTTCTTCAAAAC	5638
4663	Qy	TCAAGGAATTTATCAAACTCTGCAACACTTGGGTATGAGGTAGCAAGGGAATCTCATAC	4722
5639	Db	TCAAGGAATTTATCAAACTCTGCAACACTTGGGTATGAGGTAGCAAGGGAATCTCATAC	5698
4723	Qy	TTGAAGTCGGACAGTGGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTTTATTATCAGT	4782
5699	Db	TTGAAGTCGGACAGTGGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTTTATTATCAGT	5758
4783	Qy	GAGTCAGTCATCAGAGATCTCTACCGCGGACGGCATCGTGGCGGACCTGACAGGTCGCA	4842
5759	Db	GAGTCAGTCATCAGAGATCTCTACCGCGGACGGCATCGTGGCGGACCTGACAGGTCGCA	5818
4843	Qy	TCACCGCGGACAGGTCGGTTCGTGGCGCTATATCGCGGACATCAACCATGGGGAAG	4902
5819	Db	TCACCGCGGACAGGTCGGTTCGTGGCGCTATATCGCGGACATCAACCATGGGGAAG	5878

QY 4903 ATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTCGGCGTGGGTATGTTGCGAGGCC 4962  
DB 5879 ATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTCGGCGTGGGTATGTTGCGAGGCC 5938  
QY 4963 CCGTGGCCGGGGGACTGTGGGGCCCATCTCTCTTGCATGAC 5004  
DB 5939 CCGTGGCCGGGGGACTGTGGGGCCCATCTCTCTTGCATGAC 5998  
QY 5005 5004  
DB 5999 CCACGTTGTCTCAAAATCTCTGATGTACATTGCAACAGATAAAATATATCATCATG 6058  
QY 5005 5004  
DB 6059 AACATAAAACTGTCTGCTTACATAAACAGTAATACAGGGGTGTATGAGCCATATCA 6118  
QY 5005 5004  
DB 6119 ACGGGAACGCTTGTCTCAAGCCCGGATTAATTCACATGATGCTGATTTATATGG 6178  
QY 5005 5004  
DB 6179 GTATAAATGGGCTCGCGAATATGTCGGCAATCAGGTGCGACAATCTATCGATTGTATGG 6238  
QY 5005 5004  
DB 6239 GAAGCCGATGCGCAGAGTTGTTTCTGAACATGCGCAAGGTAGCGTTCGCAATGATGT 6298  
QY 5005 5004  
DB 6299 TACAGATGAGATGTCAGACTAACTGCTGACCGAATTTATGCTCTTCCGACCATCAA 6358  
QY 5005 5004  
DB 6359 GCATTTATCGGTACTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6418  
QY 5005 5004  
DB 6419 AGCATTCAGGTATAGAGAAATATCTGATTCAGGTGAAATATTTGTTGATCGCTGGC 6478  
QY 5005 5004  
DB 6479 AGTGTTCCTGCGCGTTGCAATCGAATTCCTGTTTGTAAATGTCCTTTTAAACAGCGATCG 6538  
QY 5005 5004  
DB 6539 CGTATTCGTCTCGCTCAGCGCAATCAGCAATGAATAACGGTTTGTGATCGGAGTGA 6598  
QY 5005 5004  
DB 6599 TTTTATGACGAGCGTAATGCTGCTGCTGTTGACAAAGTCTGGAAGAAATGCAATAGCT 6658  
QY 5005 5004  
DB 6659 TTTGCCATTCACCGGATTCAGTCGTCACTCATGTTGATTTCTCACTTGATAACCTTAT 6718  
QY 5005 5004  
DB 6719 TTTTACGAGGGGAAATTAATAGTTTGTATGATGTTGGACGAGTCGGAATCGCAGACCG 6778  
QY 5005 5004  
DB 6779 ATACCAGGATCTTCCATCTTATGGAATCTGATTAATCTGATTAATCTGATTAATCTT 6838  
QY 5005 5004  
DB 6839 ACGGCTTTTCAAAATATGTTATGATTAATCTGATTAATCTGATTAATCTGATTAATCTT 6898  
QY 5005 5004  
DB 6899 GATGCTGATGATTTTCTTAATCAGAAATGGTTAAATGTTGTTGTTAAACATCGCAGAGCAT 6958

QY 5005 5004  
DB 6959 TACGCTGACTTGACGGGACGGCGCTTGTGAATAAATCGAACTTTTGTCTGAGTTGAAG 7018  
QY 5005 5004  
DB 7019 GATCAGATCACGCACTTTCCTCCGACAAACGACACCGTTCCTGGGCAAGAAAGTTCAA 7078  
QY 5005 5004  
DB 7079 ATCACCAACTGCTCCACCTACAAACAAAGCTCTCATCAACCGTGGCTCCTCCTCCTCTG 7138  
QY 5005 5004  
DB 7139 CTGSGATGATGGGGGANTTCAGCCCTGATAGTCAGCAACACCTTCTTACGAGGCGAGA 7198  
QY 5005 5042  
DB 7199 CCTCAGCCGCCCCCTCCAGGTCCTCCAGCGCGCTGCTCAACGCGCTCAACCTA 7258  
QY 5043 CTACTGGGCTGCTTCTTAATGAGAGTCGCATTAAGGAGAGCGTCGAGTATCTATGAT 5102  
DB 7259 CTACTGGGCTGCTTCTTAATGAGAGTCGCATTAAGGAGAGCGTCGAGTATCTATGAT 7318  
QY 5103 GGAAGTATGGGAATGGTATACCCGCAATTCCTCAGTGTCTTGAAGTCTCTCTATCAGATTA 5162  
DB 7319 GGAAGTATGGGAATGGTATACCCGCAATTCCTCAGTGTCTTGAAGTCTCTCTATCAGATTA 7378  
QY 5163 TGCCCAACTAAAGCAACCGGAGGAGGATTCATGTAATTTCTCTGACTTTGTTGCTCA 5222  
DB 7379 TGCCCAACTAAAGCAACCGGAGGAGGATTCATGTAATTTCTCTGACTTTGTTGCTCA 7438  
QY 5223 TCAGTAGACTCGACTGTGAGACTATCTCGGTATGACAGCAGAAATCTCCTCTCTCGAG 5282  
DB 7439 TCAGTAGACTCGACTGTGAGACTATCTCGGTATGACAGCAGAAATCTCCTCTCTCGAG 7498  
QY 5283 ACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTGTATCAGGAACAACCTCTTGT 5342  
DB 7499 ACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTGTATCAGGAACAACCTCTTGT 7558  
QY 5343 CGAACTTTTGGTGGCTTGAACCTATAAATGATAGAGTGAATGTCGGGTAGAAATGGA 5402  
DB 7559 CGAACTTTTGGTGGCTTGAACCTATAAATGATAGAGTGAATGTCGGGTAGAAATGGA 7618  
QY 5403 GCGGCAATGCTTACCTTCTGGACCTTCAAGAGTATGTAGGTTTGTAGATACTGATG 5462  
DB 7619 GCGGCAATGCTTACCTTCTGGACCTTCAAGAGTATGTAGGTTTGTAGATACTGATG 7678  
QY 5463 CCAACTTCAGTGAACAACTGCTATTTGTTCAACCAATTCGGAATCCAGAAATCAA 5522  
DB 7679 CCAACTTCAGTGAACAACTGCTATTTGTTCAACCAATTCGGAATCCAGAAATCAA 7738  
QY 5523 GTTGTCTCTACTATTCATCCAGCCAGTCGGTCTTGAACCTGACAAATAGTGTCTCG 5582  
DB 7739 GTTGTCTCTACTATTCATCCAGCCAGTCGGTCTTGAACCTGACAAATAGTGTCTCG 7798  
QY 5583 TGTGTTGAGGTCACTTGTATGAATAAATCTAGTCTTGTATCTAAATAATCTTGTAGAG 5642  
DB 7799 TGTGTTGAGGTCACTTGTATGAATAAATCTAGTCTTGTATCTAAATAATCTTGTAGAG 7858  
QY 5643 CCAAGCCATAAATACCAAAATCTAAACTCTTTTAAACGTTTAAAGGACAACTATGTC 5702  
DB 7859 CCAAGCCATAAATACCAAAATCTAAACTCTTTTAAACGTTTAAAGGACAACTATGTC 7918  
QY 5703 TGCCCTGATTTAAACCCCAATCAGCTCGTGTCTGATCTCTCATCAACTTGAGGGGCACTA 5762  
DB 7919 TGCCCTGATTTAAACCCCAATCAGCTCGTGTCTGATCTCTCATCAACTTGAGGGGCACTA 7978  
QY 5763 TCTGTTTATAGAAATTTGCGGAGATCGGATATCGAGAAAGGTACGCTGATTTTAA 5822  
DB 7979 TCTGTTTATAGAAATTTGCGGAGATCGGATATCGAGAAAGGTACGCTGATTTTAA 8038  
QY 5823 CGTGAATTTTATCTCAAGATCTCTGCTCGCGCGCTTTCGGTGTAGTACGGTGAACCTCT 5882

Db 8039 CGTGAATTTATCTCAAGATCTCTGCTCGCGGTTTCGGTGATGACGGTGAAACCTCT 8098  
QY 5883 GACATGAGCTCCCGGAGACGCTACAGCTTGCTGTAAAGCGATGCCGGAGCAGAC 5942  
Db 8099 GACATGAGCTCCCGGAGACGCTACAGCTTGCTGTAAAGCGATGCCGGAGCAGAC 8158  
QY 5943 AAGCCCGTCAGGCGCGCTCAGCGGCTGTTGGGGGTGTCGGGGCGCAGCCATGCCAGT 6002  
Db 8159 AAGCCCGTCAGGCGCGCTCAGCGGCTGTTGGGGGTGTCGGGGCGCAGCCATGCCAGT 8218  
QY 6003 CACGTAGCGATAGCGGAGTGATATCTGCTTAACTATCGGCATCAGAGCAGATTGTACT 6062  
Db 8219 CACGTAGCGATAGCGGAGTGATATCTGCTTAACTATCGGCATCAGAGCAGATTGTACT 8278  
QY 6063 GAGAGTCACCATATGCGGCTGTAATACCGCACAGATCGCTAGGAGAAATACCGCAT 6122  
Db 8279 GAGAGTCACCATATGCGGCTGTAATACCGCACAGATCGCTAGGAGAAATACCGCAT 8338  
QY 6123 CAGGCGCTCTTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 6182  
Db 8339 CAGGCGCTCTTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 8398  
QY 6183 AGCGGTATCAGCTTACTCAAGCGGTAATACGGTTATCCAGATCAGGGGATACGC 6242  
Db 8399 AGCGGTATCAGCTTACTCAAGCGGTAATACGGTTATCCAGATCAGGGGATACGC 8458  
QY 6243 AGGAAAGAACATGTCAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAAAGGCCGTT 6302  
Db 8459 AGGAAAGAACATGTCAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAAAGGCCGTT 8518  
QY 6303 GTCGCGCTTTTCCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCAOACCTCAAG 6362  
Db 8519 GTCGCGCTTTTCCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCAOACCTCAAG 8578  
QY 6363 TCAGAGGTGGGAAACCCGACGAGACTATAAGATACCAGCGGTTCCCGCTGGAAGCTC 6422  
Db 8579 TCAGAGGTGGGAAACCCGACGAGACTATAAGATACCAGCGGTTCCCGCTGGAAGCTC 8638  
QY 6423 CCGTGGCGCTCTCTGTTCCGACCCCTGCGGTTACCGGATACCTGTCCGCTTCTCCC 6482  
Db 8639 CCGTGGCGCTCTCTGTTCCGACCCCTGCGGTTACCGGATACCTGTCCGCTTCTCCC 8698  
QY 6483 TTCGGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGT 6542  
Db 8699 TTCGGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGT 8758  
QY 6543 CGTTCGCTCCAGCTGGCTGTGTCAGCAACCCCGCTTCAGCCGACCGCTCGCCTT 6602  
Db 8759 CGTTCGCTCCAGCTGGCTGTGTCAGCAACCCCGCTTCAGCCGACCGCTCGCCTT 8818  
QY 6603 ATCCGGTAACATCTCTTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGC 6662  
Db 8819 ATCCGGTAACATCTCTTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGC 8878  
QY 6663 AGCCACTGGTAACAGGATTCAGAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAA 6722  
Db 8879 AGCCACTGGTAACAGGATTCAGAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAA 8938  
QY 6723 GTGTGGCTTAACTACGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGCTGAA 6782  
Db 8939 GTGTGGCTTAACTACGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGCTGAA 8998  
QY 6783 GCCAGTTACCTTCGAAAGAGTTGAGTCTCTTTGATCCGGCAAAACAAACCCGCTGG 6842  
Db 8999 GCCAGTTACCTTCGAAAGAGTTGAGTCTCTTTGATCCGGCAAAACAAACCCGCTGG 9058  
QY 6843 TAGCGGTGGTTTTTTTGTGTAACGACGAGTACCGGCGAAGAAAGGATCTCAAGA 6902  
Db 9059 TAGCGGTGGTTTTTTTGTGTAACGACGAGTATACCGGCGAAGAAAGGATCTCAAGA 9118  
QY 6903 AGATCCTTTGATCTTTTCTACGGGCTTGACGCTCAGTGGAAACGAAACTCACGTTAAG 6962

Db 9119 AGATCCTTTGATCTTTTCTACGGGCTTGACGCTCAGTGAACGAAACCTCAGTTAAGG 9178  
QY 5963 GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAAAATG 7022  
Db 9179 GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAAAATG 9238  
QY 7023 AAGTTTTAAATCAATCTAAAGTATATGAGTAACCTTGGTCTGACAGTTTACCAATGCTT 7082  
Db 9239 AAGTTTTAAATCAATCTAAAGTATATGAGTAACCTTGGTCTGACAGTTTACCAATGCTT 9298  
QY 7083 AATCAGTAGGACCTATCTCAGCGATCTGTCTATTTTGGTTCATCATAGTTGCTGACT 7142  
Db 9299 AATCAGTAGGACCTATCTCAGCGATCTGTCTATTTTGGTTCATCATAGTTGCTGACT 9358  
QY 7143 CCCGCTCGTGTAGATAACTACGATACGAGGCTTACCATCTGGCCCCAGTGTCTCAAT 7202  
Db 9359 CCCGCTCGTGTAGATAACTACGATACGAGGCTTACCATCTGGCCCCAGTGTCTCAAT 9418  
QY 7203 GATACCGGAGACCCAGCTCCAGGCTCCAGATTTATCAGCAATTAACCCAGCCCGG 7262  
Db 9419 GATACCGGAGACCCAGCTCCAGGCTCCAGATTTATCAGCAATTAACCCAGCCCGG 9478  
QY 7263 AAGGCGGAGCGCAGAGTGGTCTCGAACTTTATCCGCTCCATCCAGTCTATTAAATG 7322  
Db 9479 AAGGCGGAGCGCAGAGTGGTCTCGAACTTTATCCGCTCCATCCAGTCTATTAAATG 9538  
QY 7323 TTGCGCGGAAGCTAGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCAACGTTTGGCCAT 7382  
Db 9539 TTGCGCGGAAGCTAGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCAACGTTTGGCCAT 9598  
QY 7383 TGCTGAGGACATCGTGGTGTACGCTCGCTGCTTGGTATGGCTTCAATCAGCTCCGGTTC 7442  
Db 9599 TGCTGAGGACATCGTGGTGTACGCTCGCTGCTTGGTATGGCTTCAATCAGCTCCGGTTC 9658  
QY 7443 CCAACGATCAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGTCCCTT 7502  
Db 9659 CCAACGATCAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGTCCCTT 9718  
QY 7503 CGTCTCTCCGATCTGTGAGAAAGTTGCGCGCAGTGTATCTACTCATGTTATGCG 7562  
Db 9719 CGTCTCTCCGATCTGTGAGAAAGTTGCGCGCAGTGTATCTACTCATGTTATGCG 9778  
QY 7563 AGCACTGCAATCTCTTACTGTCTATGCCATCCGTAAGATGCTTTCTGCTGACTGCTGA 7622  
Db 9779 AGCACTGCAATCTCTTACTGTCTATGCCATCCGTAAGATGCTTTCTGCTGACTGCTGA 9838  
QY 7623 GTACTCAACCAAGTCAATCTGAGATAGTGTATGCGCGACCGAGTTGCTCTTGCCCGGC 7682  
Db 9839 GTACTCAACCAAGTCAATCTGAGATAGTGTATGCGCGACCGAGTTGCTCTTGCCCGGC 9898  
QY 7683 GTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGGA 7742  
Db 9899 GTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGGA 9958  
QY 7743 ACGTTCTTGGGGCGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA 7802  
Db 9959 ACGTTCTTGGGGCGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA 10018  
QY 7803 ACCCACTCGTGCAACCACTGTCTTCAAGTCTTTTACCTTTTACCAAGGTTCTGGTG 7862  
Db 10019 ACCCACTCGTGCAACCACTGTCTTCAAGTCTTTTACCTTTTACCAAGGTTCTGGTG 10078  
QY 7863 AGCAAAAACAGGAAGCAAAATCCCGCAAAAAGGGAATAAGGGCCACACGGAATGTTG 7922  
Db 10079 AGCAAAAACAGGAAGCAAAATCCCGCAAAAAGGGAATAAGGGCCACACGGAATGTTG 10138  
QY 7923 AATACTCATCTCTCTCTTTTCAATATTTAATTTGAAGCAATTTATCAGGTTATTTGCTCAT 7982  
Db 10139 AATACTCATCTCTCTCTTTTCAATATTTAATTTGAAGCAATTTATCAGGTTATTTGCTCAT 10198  
QY 7983 GACCGATACATTTTGAATGATTTAGAAAAATAAACAANAAGGGTTCCCGGCACTT 8042  
Db 10199 GACCGATACATTTTGAATGATTTAGAAAAATAAACAANAAGGGTTCCCGGCACTT 10258

8043 TCCCGGAAAGTGCACCTGACGCTCTAAGAAACCATTTATCATGACATTAACCTATAA 8102  
 Db TCCCGGAAAGTGCACCTGACGCTCTAAGAAACCATTTATCATGACATTAACCTATAA 10318  
 8103 AAATAGGCGTATCAGAGGCGCTTCGCTCTCAAGATTAATTTCTCATGTTTCAGAGCTT 8162  
 Db AAATAGGCGTATCAGAGGCGCTTCGCTCTCAAGATTAATTTCTCATGTTTCAGAGCTT 10378  
 8163 ATCATCGATAGCTGACTCATCTGTTGGTATTGTTGAATAGAGCGAGATCGGGAAACACTGAA 8222  
 Db ATCATCGATAGCTGACTCATCTGTTGGTATTGTTGAATAGAGCGAGATCGGGAAACACTGAA 10438  
 8223 AAATAACAGTTATTATTTCG 8241  
 Db AAATAACAGTTATTATTTCG 10457  
 RESULT 13  
 AX417109 10462 bp DNA linear PAT 14-JUN-2002  
 LOCUS Sequence 34 from Patent WO0198344.  
 DEFINITION  
 ACCESSION AX417109  
 VERSION AX417109.1 GI:21449696  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 unclassified.  
 1  
 REFERENCE  
 AUTHORS Ling, L. E. and Sanicola-Madel, M.  
 TITLE Angiogenesis-modulating compositions and uses  
 JOURNAL Patent: WO 0198344-A 34 27-DEC-2001;  
 BIOGEN, INC. (US)  
 FEATURES  
 source Location/Qualifiers  
 1..10462  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"  
 /note="Plasmid pUB116"  
 ORIGIN  
 Query Match 69.3%; Score 5709.8; DB 6; Length 10462;  
 Best Local Similarity 77.4%; Pred. No. 0;  
 Matches 8096; Conservative 0; Mismatches 142; Indels 2221; Gaps 8;  
 2 GATCTAACATCAAGAGCGAAGGTTGAATCAAAACCTTTTGGCCATCCGACATCCACAGG 61  
 Db GATCTAACATCAAGAGCGAAGGTTGAATCAAAACCTTTTGGCCATCCGACATCCACAGG 60  
 62 TCCATTCTCACATAAGTGCACCAACGCAACAGAGGGGATACACTAGCAGCAGACCGTT 121  
 Db TCCATTCTCACATAAGTGCACCAACGCAACAGAGGGGATACACTAGCAGCAGACCGTT 120  
 122 GCAACGAGGAGACCTCCACTCTCTCTCTCAACACCCACTTTTGGCCATCGAAACCA 181  
 Db GCAACGAGGAGACCTCCACTCTCTCTCTCAACACCCACTTTTGGCCATCGAAACCA 180  
 182 GCCAGTTATTGGGCTTGATTGGAGCTCGCTCAITCCAAATTCCTTATTAGGCTACTAA 241  
 Db GCCAGTTATTGGGCTTGATTGGAGCTCGCTCAITCCAAATTCCTTATTAGGCTACTAA 240  
 242 CACCATGACTTTATTAGCTGTCTATCTGCGCCCTTGGCGAGGTCATGTTGTTTAT 301  
 Db CACCATGACTTTATTAGCTGTCTATCTGCGCCCTTGGCGAGGTCATGTTGTTTAT 300  
 302 TTCGAAATGCAACAGCTCCGATTAACACCGGAAACATCACTCCAGATGAGGCTTTCTGA 361  
 Db TTCGAAATGCAACAGCTCCGATTAACACCGGAAACATCACTCCAGATGAGGCTTTCTGA 360  
 362 GTGTGGGTCAAAATAGTTTCATGTTCCCGAAATGGCCCAAACTGACAGTTTAAACGCTG 421  
 Db GTGTGGGTCAAAATAGTTTCATGTTCCCGAAATGGCCCAAACTGACAGTTTAAACGCTG 420

422 TCTTGGACCTAATATGACAAAAGCGTGATCTCATCAAGATGAACCTAAGTTGGTTCTG 481  
 Db TCTTGGACCTAATATGACAAAAGCGTGATCTCATCAAGATGAACCTAAGTTGGTTCTG 480  
 482 TGAATGTAAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTGGCCATACCGTTTGTG 541  
 Db TGAATGTAAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTGGCCATACCGTTTGTG 540  
 542 TTGTTTGTATGATGAGCAATGCTCAAAATTAATCTCATTAATGCTTGGCGAGTCTC 601  
 Db TTGTTTGTATGATGAGCAATGCTCAAAATTAATCTCATTAATGCTTGGCGAGTCTC 600  
 602 TCTATCGCTTCTGAACCCCGGTGACCTGTGCCGAAACGCAAAATGGGAAACACCCGCTT 661  
 Db TCTATCGCTTCTGAACCCCGGTGACCTGTGCCGAAACGCAAAATGGGAAACACCCGCTT 660  
 662 TTTGGATGATTATGCAATGCTCCACATTTGATGCTTCCAGATTTGGTGGGAATCTG 721  
 Db TTTGGATGATTATGCAATGCTCCACATTTGATGCTTCCAGATTTGGTGGGAATCTG 720  
 722 CTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTTGACAGCAATA 781  
 Db CTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTTGACAGCAATA 780  
 782 TATAACAGAGGAAAGTGGCCCTGTCTTAAACCTTTTATTTATCATCAATTTAGCTTA 841  
 Db TATAACAGAGGAAAGTGGCCCTGTCTTAAACCTTTTATTTATCATCAATTTAGCTTA 840  
 842 CTTTCATTAATTCGGACTGCTTCCAAATTCACAAGCTTTTGAATTTAAACGACTTTTAAAGAC 901  
 Db CTTTCATTAATTCGGACTGCTTCCAAATTCACAAGCTTTTGAATTTAAACGACTTTTAAAGAC 900  
 902 AACTTTGAGAAGATCAAAAACAACTTAATTTTGAAGGATCCAAACGATGAGATTTCTT 961  
 Db AACTTTGAGAAGATCAAAAACAACTTAATTTTGAAGGATCCAAACGATGAGATTTCTT 960  
 962 CAAATTTTACTGCACTTTTATTCGAGCATCTCCGATAGCTGCTCCAGTCAACACTA 1021  
 Db CAAATTTTACTGCACTTTTATTCGAGCATCTCCGATAGCTGCTCCAGTCAACACTA 1020  
 1022 CAACAGAGATGAAGCGCACAAATTCGGCTGAAGCTGTCTCATCGTTTACTTCAGATTAG 1081  
 Db CAACAGAGATGAAGCGCACAAATTCGGCTGAAGCTGTCTCATCGTTTACTTCAGATTAG 1080  
 1082 AAGGGATTTTCGATGTTGCTGTTTCCATTTTCCACAGCAGCAAAATACCGGTTATTGT 1141  
 Db AAGGGATTTTCGATGTTGCTGTTTCCATTTTCCACAGCAGCAAAATACCGGTTATTGT 1140  
 1142 TTATAAATACTACTATTGTCAGCATTTGCTGTAAAGAGAGAGGGGTATCTCTCGAGAAA 1201  
 Db TTATAAATACTACTATTGTCAGCATTTGCTGTAAAGAGAGAGGGGTATCTCTCGAGAAA 1200  
 1202 GAGAGG----- 1207  
 Db GAGAGG----- 1207  
 1201 GATCGGACCGGGAGGGGTTGGGAAAGAGGAGGACCCCAAAAGCTGACCCCTTTAG 1260  
 Db GATCGGACCGGGAGGGGTTGGGAAAGAGGAGGACCCCAAAAGCTGACCCCTTTAG 1259  
 1261 CTTAAGCAGTTTATCCCAATGTGGCGGAGAGACCCCTAGGCCAGCGGAGGTATG 1320  
 Db CTTAAGCAGTTTATCCCAATGTGGCGGAGAGACCCCTAGGCCAGCGGAGGTATG 1319  
 1321 AAGGGAAGATCTCCAGAAATCCGAGCGATTTAAGGAACTCACCCCAATTAACACCCCG 1380  
 Db AAGGGAAGATCTCCAGAAATCCGAGCGATTTAAGGAACTCACCCCAATTAACACCCCG 1379  
 1381 ACATCATATTTAAGGATGAAGAAACACCGGAGCGGACAGGCTGATGACTCAGAGGTGA 1440  
 Db ACATCATATTTAAGGATGAAGAAACACCGGAGCGGACAGGCTGATGACTCAGAGGTGA 1439  
 1441 AGGACAAGTTGAACGCTTTGGCCATCTCGGTGATGAACAGTGGCGGAGGAGTGAACCTGC 1500  
 Db AGGACAAGTTGAACGCTTTGGCCATCTCGGTGATGAACAGTGGCGGAGGAGTGAACCTGC 1499



QY 2744 ATCTGGATCAGCGTATTTATCAGCAATAAATAAGAACTTTCAGAAAGGCCAGCAGGCATGTC 2803  
 Db 3720 ATCTGGATCAGCGTATTTATCAGCAATAAATAAGAACTTTCAGAAAGGCCAGCAGGCATGTC 3779  
 QY 2804 AATACTACAGAGGCTGATGTCATTTTGAACCATCATCTTGGCAGCAGTACGAACTG 2863  
 Db 3780 AATACTACAGAGGCTGATGTCATTTTGAACCATCATCTTGGCAGCAGTACGAACTG 3839  
 QY 2864 GTTTCCTGGACCAATAATTTTGTACACTTAGGAACAGATTTCTGTCGTAAGCCATAGC 2923  
 Db 3840 GTTTCCTGGACCAATAATTTTGTACACTTAGGAACAGATTTCTGTCGTAAGCCATAGC 3899  
 QY 2924 AGCTACTGCTGGCGGCTCTGCTAGCAGATACATTTAGCAGCAATACCTTTGTTGGGCAAC 2983  
 Db 3900 AGCTACTGCTGGCGGCTCTGCTAGCAGATACATTTAGCAGCAATACCTTTGTTGGGCAAC 3959  
 QY 2984 GTAGATGACTTCTGGGTAAAGGTACCATCTTCTTAGGTGGAGTGCATAAAACAATTTTC 3043  
 Db 3960 GTAGATGACTTCTGGGTAAAGGTACCATCTTCTTAGGTGGAGTGCATAAAACAATTTTC 4019  
 QY 3044 TTTGCAACACGACAACTTTGGCAGGAACAACCCAGCATCAGGGAAGTGGAAGGCGAATTCG 3103  
 Db 4020 TTTGCAACACGACAACTTTGGCAGGAACAACCCAGCATCAGGGAAGTGGAAGGCGAATTCG 4079  
 QY 3104 GGTTCACAGCAATATAGAGGCCAACTTTCTCAATAGGTCTTCGCAAAACGAGAGCAGAC 3163  
 Db 4080 GGTTCACAGCAATATAGAGGCCAACTTTCTCAATAGGTCTTCGCAAAACGAGAGCAGAC 4139  
 QY 3164 TACACCGGCGAAGTCTCAACTTGCAGAGTCTCGGTAGTTAGCTTTCATGGAATTTCCCT 3223  
 Db 4140 TACACCGGCGAAGTCTCAACTTGCAGAGTCTCGGTAGTTAGCTTTCATGGAATTTCCCT 4199  
 QY 3224 GACGTATCTATAGAGAGATCAATGGCTCTCTTAAGCTTATCTGGCAATTCGATAGTTTC 3283  
 Db 4200 GACGTATCTATAGAGAGATCAATGGCTCTCTTAAGCTTATCTGGCAATTCGATAGTTTC 4259  
 QY 3284 CTCTGGGAAAGAGGCTTCTAACAAGAGTGTCTTCAAGCGACTCCATCAAACTTGGCAGT 3343  
 Db 4260 CTCTGGGAAAGAGGCTTCTAACAAGAGTGTCTTCAAGCGACTCCATCAAACTTGGCAGT 4319  
 QY 3344 TAGTCTTAAAAGGCTTTGTCCACATTTTGAAGCAATTTGCGCAATTTGGTTGACTAA 3403  
 Db 4320 TAGTCTTAAAAGGCTTTGTCCACATTTTGAAGCAATTTGCGCAATTTGGTTGACTAA 4379  
 QY 3404 TTCCATATCTGTCGCTTTCTGATAGGAGCAGGAGGCACTTCAATTTCTTGTA 3463  
 Db 4380 TTCCATATCTGTCGCTTTCTGATAGGAGCAGGAGGCACTTCAATTTCTTGTA 4439  
 QY 3464 GGAGGCTTTAGAAAGCTCAATTTTGCACAAATTTCAATACGACCTTTCAGAAAGGACTTCTTT 3523  
 Db 4440 GGAGGCTTTAGAAAGCTCAATTTTGCACAAATTTCAATACGACCTTTCAGAAAGGACTTCTTT 4499  
 QY 3524 AGTTTGGATTTCTTTAGGTTGTTCTTGGTATCTGCTGGCTTGGCATCTCCCTTCCCT 3583  
 Db 4500 AGTTTGGATTTCTTTAGGTTGTTCTTGGTATCTGCTGGCTTGGCATCTCCCTTCCCT 4559  
 QY 3584 TCTAGTGACCTTTAGGCACTTCAATCCAGGTTTCTCTCCACCTCTGCTCCAAACGTCACACC 3643  
 Db 4560 TCTAGTGACCTTTAGGCACTTCAATCCAGGTTTCTCTCCACCTCTGCTCCAAACGTCACACC 4619  
 QY 3644 GTACTTGGCACTATCACTAATGCAAAATAAATAAGTTCAGCAGCAATTTCCCGAGGTATATC 3703  
 Db 4620 GTACTTGGCACTATCACTAATGCAAAATAAATAAGTTCAGCAGCAATTTCCCGAGGTATATC 4679  
 QY 3704 TTCCCTTGGATTTAGCTTCTGCAAGTTTCAATCAGCTTCTCTCCCTAATTTTAGGTTTC-AAACA 3762  
 Db 4680 TTCCCTTGGATTTAGCTTCTGCAAGTTTCAATCAGCTTCTCTCCCTAATTTTAGGTTTCABAACA 4739  
 QY 3763 AAACTTCGTCGTCAAATTAACCGTTTGGTATTAAGAACTTCTTGGAGCAATTTGCTTCTAGCAT 3822  
 Db 4740 AAACTTCGTCGTCAAATTAACCGTTTGGTATTAAGAACTTCTTGGAGCAATTTGCTTCTAGCAT 4799

QY 3823 CCCACAGGTGGCTTCCATGGCTCTAGACCCCTTTGATTGGCCAAAACAGGAAGTGGCTT 3882  
 Db 4800 CCCACAGGT-GCTTCCATGGCTCTAGACCCCTTTGATTGGCCAAAACAGGAAGTGGCTT 4858  
 QY 3883 CCAAGTGACAGAAACCAACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCAC 3942  
 Db 4859 CCAAGTGACAGAAACCAACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCAC 4918  
 QY 3943 AATCAATTCGANTACCCAGAACCTTTTTCAGTTGCTCAGATGTAGCACCCTTTATACACA 4002  
 Db 4919 AATCAATTCGANTACCCAGAACCTTTTTCAGTTGCTCAGATGTAGCACCCTTTATACACA 4978  
 QY 4003 AACCGTGACGAGATTTGGTAGCTCCAGTTTGTGTCCTTATAGCCTTCCGGAATAGACT 4062  
 Db 4979 AACCGTGACGAGATTTGGTAGCTCCAGTTTGTGTCCTTATAGCCTTCCGGAATAGACT 5038  
 QY 4063 TTTTGGACGAGTACACAGGCCCCAACGAGTAAATTTAGAGAGTTCAGCCACCAAGTAGTGA 4122  
 Db 5039 TTTTGGACGAGTACACAGGCCCCAACGAGTAAATTTAGAGAGTTCAGCCACCAAGTAGTGA 5098  
 QY 4123 ATAGACCAATCGGGCGGTTCAGTGTCAAGAGCGCAACAAAATTTTCACTGACAGGAACT 4182  
 Db 5099 ATAGACCAATCGGGCGGTTCAGTGTCAAGAGCGCAACAAAATTTTCACTGACAGGAACT 5158  
 QY 4183 TTTTGCATCTTTCAGAAAAGTTTCGTATTCAGTGTCAATTTGCGAGCATCAATAATGGGGA 4242  
 Db 5159 TTTTGCATCTTTCAGAAAAGTTTCGTATTCAGTGTCAATTTGCGAGCATCAATAATGGGGA 5218  
 QY 4243 TTATACGAGAACCAACAGTGGAGTCACTCAACACTTTTGGGTCTCAGAAAAGCAT 4302  
 Db 5219 TTATACGAGAACCAACAGTGGAGTCACTCAACACTTTTGGGTCTCAGAAAAGCAT 5278  
 QY 4303 AAAAGTTTCTACTACCGCCATTTAGTGAATTTTCAAAATCGCCAGTGGAGAAAGAAAG 4362  
 Db 5279 AAAAGTTTCTACTACCGCCATTTAGTGAATTTTCAAAATCGCCAGTGGAGAAAGAAAG 5338  
 QY 4363 GCACAGCATATACGANTTAGGGCAAGGATGCAACTTTATCAACCGGTCTCTATAGA 4422  
 Db 5339 GCACAGCATATACGANTTAGGGCAAGGATGCAACTTTATCAACCGGTCTCTATAGA 5398  
 QY 4423 TAAACCTAGCGCTGGGATCATCTTTTGGCAAACTCTTTCTGCCAAATCTAGTCCAAA 4482  
 Db 5399 TAAACCTAGCGCTGGGATCATCTTTTGGCAAACTCTTTCTGCCAAATCTAGTCCAAA 5458  
 QY 4483 TCACCTTATTGATACCATTTATGTAACACTTTGAGCAAGTTTGCATAGCTCTCTCAAT 4542  
 Db 5459 TCACCTTATTGATACCATTTATGTAACACTTTGAGCAAGTTTGCATAGCTCTCTCAAT 5518  
 QY 4543 GGTCTCTGTAAACGATGACTCAACTTCACATTAACCTTGAAGTTCAGTTCAGTGAAGTGA 4602  
 Db 5519 GGTCTCTGTAAACGATGACTCAACTTCACATTAACCTTGAAGTTCAGTTCAGTGAAGTGA 5578  
 QY 4603 ACTTTCATCAGGTTTGTGAGCTGGTTCAGCAGCATAGGAAACACGCGCTTTTCTACCAAC 4662  
 Db 5579 ACTTTCATCAGGTTTGTGAGCTGGTTCAGCAGCATAGGAAACACGCGCTTTTCTACCAAC 5638  
 QY 4663 TCAAGAAATATCAAACTCTGCAACACTTGGTATGAGGTAGCAAGGAAATGTCATAC 4722  
 Db 5639 TCAAGAAATATCAAACTCTGCAACACTTGGTATGAGGTAGCAAGGAAATGTCATAC 5698  
 QY 4723 TTGAAGTCGACAGTGTAGTGTCTTGAAGAAATTTCTGAAGCCGCTATTTTATTCAGT 4782  
 Db 5699 TTGAAGTCGACAGTGTAGTGTCTTGAAGAAATTTCTGAAGCCGCTATTTTATTCAGT 5758  
 QY 4783 GAGTCAGTCATCAGGAGATCTCTACCGCGGAGGAGCATCTGTGGCCGACCTGAGGTCCGCA 4842  
 Db 5759 GAGTCAGTCATCAGGAGATCTCTACCGCGGAGGAGCATCTGTGGCCGACCTGAGGTCCGCA 5818  
 QY 4843 TCACCGGCGCCACAGGTGGGTGCTGCGGCTTATATCGCGGACATCACCGATGGGGAAG 4902  
 Db 5819 TCACCGGCGCCACAGGTGGGTGCTGCGGCTTATATCGCGGACATCACCGATGGGGAAG 5878  
 QY 4903 ATCGGCTCGCCACTTCGCGGCTCATAGCGCTTTGTTTGGCGGTGGGTATGTTGGCAGGCC 4962

Db	5879	ATCGGGTTCGCCACTTCGGGCTCATGAGCGCTTGTTCGGCGTGGGTATGGTGGCAGGCC	5938
Qy	4963	CCGTGGCGGGGAGCTGTTGGGGCCCATCTCCTTGGCATGCAC	5004
Db	5939	CCGTGGCGGGGAGCTGTTGGGGCCCATCTCCTTGGACCTGCAGGGGGGGGGGNAAG	5998
Qy	5005	-----	5004
Db	5999	CCAGTGTGTCTCAAAATCTCTGATGTTACATTTGCACAGATAAAATATATCATCATG	6058
Qy	5005	-----	5004
Db	6059	AACAATAAACTGCTCTGCTTACATAAACAAGTAACAGAGGGGTGTTATGAGCCCATATCA	6118
Qy	5005	-----	5004
Db	6119	ACGGGAACGTCCTTCTCAAGGCGCGATTAAATTCGAACATGATGCTGATTATATGG	6178
Qy	5005	-----	5004
Db	6179	GTATAAATGGCTCCCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGAATTGATGG	6238
Qy	5005	-----	5004
Db	6239	GAAGCCCGATGGCCAGAGTTGTTCTGAACATGCGCAAGGTAGCGTGGCAATGATGT	6298
Qy	5005	-----	5004
Db	6299	TACAGATGAGATGGTTCAGACTAAACTGGCTGACGGAAATTTATGGCTCTCCGACCATCAA	6358
Qy	5005	-----	5004
Db	6359	GCAATTTATCCGTACTCTCTGATGATGATGATGTTACTCACACTCGGATCCCGGGGAAAC	6418
Qy	5005	-----	5004
Db	6419	AGCAATCCAGGTATTAGAGAATATCCCTGATTGAGTGAAAATATTGTTGATGCGCTGGC	6478
Qy	5005	-----	5004
Db	6479	AGTGTCTGGCCGGTTCATTCGATTCCTGTTGTAATTTGTCCTTTTAAACAGCATGG	6538
Qy	5005	-----	5004
Db	6539	CGTATTTCTGCTCGCTCAGGCGCAATCACGAATGAATACCGTTTGGTTGATGCGAGTGA	6598
Qy	5005	-----	5004
Db	6599	TTTTGATGACGAGCGTAATGGCTGGCTGTTGAACAGACTCGGAAGAAATGCAATAAGCT	6658
Qy	5005	-----	5004
Db	6659	TTTGCCATTTCCACCGGATTCAGTCGTCACCTCATGGTGATTTCTCATTGATAACCTTAT	6718
Qy	5005	-----	5004
Db	6719	TTTTGACGAGGGAATTAATAGTTGATTGATGTTGGAAGTGGGAATCGCAGACCG	6778
Qy	5005	-----	5004
Db	6779	ATACCAGGATCTTGCCATCCTATGGAATGCGCTCGGTGAGTTTCTCCTTCATTACAGAA	6838
Qy	5005	-----	5004
Db	6839	ACGGCTTTTCAAAATATGGTATTGATAATCCTGATGAATAAATTCAGTTTCATT	6898
Qy	5005	-----	5004
Db	6899	GATGCTCGATGAGTTTCTTAATCAGAAATGGTTAAATGGTTGTAACACTGCGCAGAGCAT	6958
Qy	5005	-----	5004

Db	6959	TACGCTGACTTGACGGGACGGCGGCTTGTGTAATAAATCGAACTTTTGCTGAGTTGAAG	7018
Qy	5005	-----	5004
Db	7019	GATCAGATCAGCATCTTCCGCAACAACAGACCGTTCCGTGGCAAGCAAAAGTTCAAA	7078
Qy	5005	-----	5004
Db	7079	ATCACAACTGGTCCACCTACAAACAAGCTCTCATCAACCGTGGCTCCCTCACTTCTGG	7138
Qy	5005	-----	5004
Db	7139	CTGGATGATGGGGCGATTTCAGGCTGGTATGAGTCAGCAACACCTTCTTCCAGAGCAGA	7198
Qy	5005	-----	5004
Db	7199	CCTCAGCGCCCCCCCCCTCAGGTCCACGGGGGGGTGCTCAACGGCTCAACCTTA	7258
Qy	5043	CTACTGGGCTGTTCTTAATGCAAGAGTCGCATAAAGGAGAGCGTCGAGTATCTATGATT	5102
Db	7259	CTACTGGGCTGTTCTTAATGCAAGAGTCGCATAAAGGAGAGCGTCGAGTATCTATGATT	7318
Qy	5103	GGAGTATGGGAATGGTGATACCCGCAATCTTCAGTGTCTTGAGGTCTCCTATCAGATT	5162
Db	7319	GGAGTATGGGAATGGTGATACCCGCAATCTTCAGTGTCTTGAGGTCTCCTATCAGATT	7378
Qy	5163	TGCCCAACTAAAGCAACCGGAGAGGAGATTTCATGGTAAATTTCTCAGCTTTTGGTCA	5222
Db	7379	TGCCCAACTAAAGCAACCGGAGAGGAGATTTCATGGTAAATTTCTCAGCTTTTGGTCA	7438
Qy	5223	TCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTGGAG	5282
Db	7439	TCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTGGAG	7498
Qy	5283	ACAGTAAATGAAGTCCCAACCAATAAGAAATCCTTCTTATCAGGACAAACTCTTGT	5342
Db	7499	ACAGTAAATGAAGTCCCAACCAATAAGAAATCCTTCTTATCAGGACAAACTCTTGT	7558
Qy	5343	CGAACTTTTTCGGTGCCTTGAACTATAAAATGTAGAGTGGATATGTCGGGTAGGAATGA	5402
Db	7559	CGAACTTTTTCGGTGCCTTGAACTATAAAATGTAGAGTGGATATGTCGGGTAGGAATGA	7618
Qy	5403	GGGGCAAAATGCTTACCTCTGACCTTCAAGAGGTATGTAGGGTTTGTAGATCTAGTG	5462
Db	7619	GGGGCAAAATGCTTACCTCTGACCTTCAAGAGGTATGTAGGGTTTGTAGATCTAGTG	7678
Qy	5463	CCAACTTCAGTGACAAAGTGTCTTATTCGTTCAAAACCATTCGGAATCCAGAGAAATCAAA	5522
Db	7679	CCAACTTCAGTGACAAAGTGTCTTATTCGTTCAAAACCATTCGGAATCCAGAGAAATCAAA	7738
Qy	5523	GTTGTTGTCTACTATTGATCCAGCCAGTCGGCTTGGAACTGACAATAGTGTCTCG	5582
Db	7739	GTTGTTGTCTACTATTGATCCAGCCAGTCGGCTTGGAACTGACAATAGTGTCTCG	7798
Qy	5583	TGTTTGGGTGATCTTTGTATGAATAAATCTAGTCTTTGATCTAATAATCTTTGACGAG	5642
Db	7799	TGTTTGGGTGATCTTTGTATGAATAAATCTAGTCTTTGATCTAATAATCTTTGACGAG	7858
Qy	5643	CCAAGCGATAAATACCAAACTCTTTAAACCTTTTAAACGTTTAAAGGACAAAGTATGC	5702
Db	7859	CCAAGCGATAAATACCAAACTCTTTAAACCTTTTAAACGTTTAAAGGACAAAGTATGC	7918
Qy	5703	TGCTGTATTAACCCCAATCAGCTCGTAGTCTCATCTCATCACTTTGAGGGGCACTA	5762
Db	7919	TGCTGTATTAACCCCAATCAGCTCGTAGTCTCATCTCATCACTTTGAGGGGCACTA	7978
Qy	5763	TCTTGTTTTGAAGAAATTTGCGGAGATGGATATCGAGAAAAAGGTACGCTGATTTTAA	5822
Db	7979	TCTTGTTTTGAAGAAATTTGCGGAGATGGATATCGAGAAAAAGGTACGCTGATTTTAA	8038
Qy	5823	CGTGAATTTTATCTCAAGATCTCTGCTCGCGGTTTGGGTGATGACGGTGAACCTCT	5882
Db	8039	CGTGAATTTTATCTCAAGATCTCTGCTCGCGGTTTGGGTGATGACGGTGAACCTCT	8098



QY	5883	GACACATGACGCTCCGGAGACGGTTCACAGCTTCTGTGAAGCGATGCGCGGAGCAGAC	5942	6963	GATTTTGGTCATGAGATTATCAAAAGGATCTTCACTAGATCCTTTTAAATTTAAATG	7022
DB	8099	GACACATGACGCTCCGGAGACGGTTCACAGCTTCTGTGAAGCGATGCGCGGAGCAGAC	8158	9179	GATTTTGGTCATGAGATTATCAAAAGGATCTTCACTAGATCCTTTTAAATTTAAATG	9238
QY	5943	AAGCCCGTCAGGGCGGTCAGCGGGTGTGGCGGGTTCGGGGCGCAGCCATGACCCAGT	6002	7023	AGTTTTTAATCAATCTAAAGTATATAGTAAGTAACTTGGTCTGACAGTTACCAATGCTT	7082
DB	8159	AAGCCCGTCAGGGCGGTCAGCGGGTGTGGCGGGTTCGGGGCGCAGCCATGACCCAGT	8218	9239	AGTTTTTAATCAATCTAAAGTATATAGTAAGTAACTTGGTCTGACAGTTACCAATGCTT	9298
QY	6003	CACGTAGCGATAGCGGAGTGTATCTGCTTAACTATGCGGCATCAGAGCAGATTGTACT	6062	7083	AATCAGTCAGGCACCATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACT	7142
DB	8219	CACGTAGCGATAGCGGAGTGTATCTGCTTAACTATGCGGCATCAGAGCAGATTGTACT	8278	9299	AATCAGTCAGGCACCATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACT	9358
QY	6063	GAGAGTCACCATATGCGGGTGTGAATAACCGCACAGATGCGTAAGGAGAGAAATACCGCAT	6122	7143	CCCCGTGTGTAGATAACTACGATACGGAGGGCTTACCATCTGCCCGAGTGTGCAAT	7202
DB	8279	GAGAGTCACCATATGCGGGTGTGAATAACCGCACAGATGCGTAAGGAGAGAAATACCGCAT	8338	9359	CCCCGTGTGTAGATAACTACGATACGGAGGGCTTACCATCTGCCCGAGTGTGCAAT	9418
QY	6123	CAGGGCTTTCGGTTCCTCGCTCAGTACTCGCTCGGCTCGGTCGGTTCGGCTGCGGG	6182	7203	GATACCGGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACACGACGACCGG	7262
DB	8339	CAGGGCTTTCGGTTCCTCGCTCAGTACTCGCTCGGCTCGGTCGGTTCGGCTGCGGG	8398	9419	GATACCGGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACACGACGACCGG	9478
QY	6183	AGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACGC	6242	7263	AAGGCCGAGCGCAGAACTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATG	7322
DB	8399	AGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACGC	8458	9479	AAGGCCGAGCGCAGAACTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATG	9538
QY	6243	AGGAAAGAACATGTAGCAAAAGGCGCAGAAAGGCCAGGAAACCGTAAAGAGCGCGGTT	6302	7323	TTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAATAGTTTGGCAACGTTGTGCCAT	7382
DB	8459	AGGAAAGAACATGTAGCAAAAGGCGCAGAAAGGCCAGGAAACCGTAAAGAGCGCGGTT	8518	9539	TTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAATAGTTTGGCAACGTTGTGCCAT	9598
QY	6303	GCTGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAG	6362	7383	TGCTCAGGCATCTGTGTGTCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7442
DB	8519	GCTGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAG	8578	9599	TGCTCAGGCATCTGTGTGTCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	9658
QY	6363	TCAGAGTGGCGAAACCCGACAGCACTATAAAGATACAGCGGTTTCCCGCTGGAAGCTC	6422	7443	CCAAAGATCAAGCGAGTTACATGATCCCCCATCTGTGTGCAAAAAGCGTTACTCCTT	7502
DB	8579	TCAGAGTGGCGAAACCCGACAGCACTATAAAGATACAGCGGTTTCCCGCTGGAAGCTC	8638	9659	CCAAAGATCAAGCGAGTTACATGATCCCCCATCTGTGTGCAAAAAGCGTTACTCCTT	9718
QY	6423	CCTGTGGCTTCTCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCCCTTCTCCC	6482	7503	CGGTCTCCGATCTGTGTGTCAGAAAGTGTGCGCGCAGTGTATCACTCATGTTATGCG	7562
DB	8639	CCTGTGGCTTCTCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCCCTTCTCCC	8698	9719	CGGTCTCCGATCTGTGTGTCAGAAAGTGTGCGCGCAGTGTATCACTCATGTTATGCG	9778
QY	6483	TTCCGGAAGCGTGGCGCTTCTCATAGCTCACCGCTGTAGTATCTCAGTTTCGGTGTAGT	6542	7563	AGCATGTGATATCTCTTCTTCTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGA	7622
DB	8699	TTCCGGAAGCGTGGCGCTTCTCATAGCTCACCGCTGTAGTATCTCAGTTTCGGTGTAGT	8758	9779	AGCATGTGATATCTCTTCTTCTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGA	9838
QY	6543	CGTTGGCTCCAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGCCGACCGCTGCGCTT	6602	7623	GTACTCAACCAAGTCACTTCTGAGTAAGTGTATGCGGCGCAGTGTCTCTTCCCGCGC	7682
DB	8759	CGTTGGCTCCAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGCCGACCGCTGCGCTT	8818	9839	GTACTCAACCAAGTCACTTCTGAGTAAGTGTATGCGGCGCAGTGTCTCTTCCCGCGC	9898
QY	6603	ATCCGGTAACCTATCGTCTTGAAGTCCAAACCGGTAAGACACGACTTATCGCCACTGCGAGC	6662	7683	GTCAAACGCGGATATAACCGGCGCAGATAGCAGAACTTTAAAGTGTCTCATTTGNA	7742
DB	8819	ATCCGGTAACCTATCGTCTTGAAGTCCAAACCGGTAAGACACGACTTATCGCCACTGCGAGC	8878	9899	GTCAAACGCGGATATAACCGGCGCAGATAGCAGAACTTTAAAGTGTCTCATTTGNA	9958
QY	6663	AGCCACTGGTAACAGGATACAGAGCGAGTATGTAGGCGGTGCTACAGATTCTTGAA	6722	7743	ACGTTCTTCCGGGCGGAAACTCTCAAGGATCTTACCGCTGTGTGATCCAGTTCGATGTA	7802
DB	8879	AGCCACTGGTAACAGGATACAGAGCGAGTATGTAGGCGGTGCTACAGATTCTTGAA	8938	9959	ACGTTCTTCCGGGCGGAAACTCTCAAGGATCTTACCGCTGTGTGATCCAGTTCGATGTA	10018
QY	6723	GTGGTGGCTTAACGCTACGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAA	6782	7803	ACCCACTCGTGCACCAACTGATCTTACGATCTTTTACTTTTCCAGCGTTTCTGGGTG	7862
DB	8939	GTGGTGGCTTAACGCTACGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAA	8998	10019	ACCCACTCGTGCACCAACTGATCTTACGATCTTTTACTTTTCCAGCGTTTCTGGGTG	10078
QY	6783	GCCAGTTACCTTCGGAAGAGTGTGTAGTCTTGTATCCGGCAAAACAAACCCGCTGG	6842	7863	ACCAAAACGAGGAGGAAATGCGGCAAAAGGATAGGGCGACACGGAATGTTG	7922
DB	8999	GCCAGTTACCTTCGGAAGAGTGTGTAGTCTTGTATCCGGCAAAACAAACCCGCTGG	9058	10079	ACCAAAACGAGGAGGAAATGCGGCAAAAGGATAGGGCGACACGGAATGTTG	10138
QY	6843	TAGCGGTGGTTTTTTTGTTCAGACGACAGATTACCGGAGAAAAGGATCTCAAGA	6902	7923	AATACTCATCTCTTCTCTTTTCAATATTTATTAAGCACTTTATCAGGTTATTTGCTCAT	7982
DB	9059	TAGCGGTGGTTTTTTTGTTCAGACGACAGATTACCGGAGAAAAGGATCTCAAGA	9118	10139	AATACTCATCTCTTCTCTTTTCAATATTTATTAAGCACTTTATCAGGTTATTTGCTCAT	10198
QY	6903	AGATCCTTTGATCTTTCTACGGGCTGTGACGCTCAGTGGAAACGAAACCTCAGTTAAGG	6962	7983	GAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGTTCCGCGCACAT	8042
DB	9119	AGATCCTTTGATCTTTCTACGGGCTGTGACGCTCAGTGGAAACGAAACCTCAGTTAAGG	9178	10199	GAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGTTCCGCGCACAT	10258
				8043	TCCCCGAAAAGTGCCACCTGACGCTCTAAGAAACCACTTATTTATCATGACATTAACCTATAA	8102

Db 10259 TCCCGAAAGTGCACCTGACGCTTAAGAAACCAATATATATCATGACATTAACCTATAA 10318  
 QY 8103 AAATAGGCGTATCAGAGGCCCTTTTCGCTTTCAGAGAAATTAATCTCATGTTTGACAGCTT 8162  
 Db 10319 AAATAGGCGTATCAGAGGCCCTTTTCGCTTTCAGAGAAATTAATCTCATGTTTGACAGCTT 10378  
 QY 8163 ATCATCGATAGCTGACTCATGTTGGTATTGTTGAATAGAGCGAGATCGGGAACACTGAA 8222  
 Db 10379 ATCATCGATAGCTGACTCATGTTGGTATTGTTGAATAGAGCGAGATCGGGAACACTGAA 10438  
 QY 8223 AAATAACAGTTATTATTTCG 8241  
 Db 10439 AAATAACAGTTATTATTTCG 10457  
 RESULT 14  
 LOCUS AX146619 10492 bp DNA linear PAT 31-MAY-2001  
 DEFINITION Sequence 81 from Patent WO0134654.  
 ACCESSION AX146619  
 VERSION AX146619.1 GI:14285012  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Strauch, K.  
 TITLE Hedgehog fusion proteins and uses  
 JOURNAL Patent: WO 0134654-A 81 17-MAY-2001;  
 BIOGEN, INC. (US)  
 FEATURES  
 source  
 1..10492  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 68.9%; Score 5677.8; DB 6; Length 10492;  
 Best Local Similarity 77.2%; Pred. No. 0;  
 Matches 8101; Conservative 0; Mismatches 137; Indels 2251; Gaps 9;  
 QY 2 GATTAACATCCAAAGAGGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 61  
 Db 1 GATTAACATCCAAAGAGGAAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 60  
 QY 62 TCCATTCTCACATTAAGTCCCAACGACAGAGGGGGATACACTAGCAGAGCCGTT 121  
 Db 61 TCCATTCTCACATTAAGTCCCAACGACAGAGGGGGATACACTAGCAGAGCCGTT 120  
 QY 122 GCAAGCGAGGAGCTCCACTCTCTCTCTCAACACCCACCTTTTGGCCATCGAAAAACCA 181  
 Db 121 GCAAGCGAGGAGCTCCACTCTCTCTCTCAACACCCACCTTTTGGCCATCGAAAAACCA 180  
 QY 182 GCCAGTTATGGGCTTGATGAGCTCGCTCATTCCTCAATTCCTTCTATTAGGCTACTAA 241  
 Db 181 GCCAGTTATGGGCTTGATGAGCTCGCTCATTCCTCAATTCCTTCTATTAGGCTACTAA 240  
 QY 242 CACCATGACTTTATTAGGCTGCTATCTTGGCCCTCGGAGGTTTCATGTTTCTTTAT 301  
 Db 241 CACCATGACTTTATTAGGCTGCTATCTTGGCCCTCGGAGGTTTCATGTTTCTTTAT 300  
 QY 302 TTCGAATGCAACAACTCGGATTAACCCGAAACATCACTCCAGATGAGGCTTTCTGA 361  
 Db 301 TTCGAATGCAACAACTCGGATTAACCCGAAACATCACTCCAGATGAGGCTTTCTGA 360  
 QY 362 GTGTGGGTCGAATAGTTTCATGTTCCCAATGGCCCAAACTGCAGTTTAAACGCTG 421  
 Db 361 GTGTGGGTCGAATAGTTTCATGTTCCCAATGGCCCAAACTGCAGTTTAAACGCTG 420  
 QY 422 TCTTGGAACTTAATATCAAAAACGCTGATCTCATCCAAAGATCAACTAAGTTTGGTTGCT 481

Db 421 TCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAGATGAATAAGTTTGGTTGCT 480  
 QY 482 TGAATGCTAAAGCGCCAGTTGTCAAAAGAAACTTCCAAAAGTCCCATACCGTTTGC 541  
 Db 481 TGAATGCTAAAGCGCCAGTTGTCAAAAGAAACTTCCAAAAGTCCCATACCGTTTGC 540  
 QY 542 TTGTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 601  
 Db 541 TTGTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 600  
 QY 602 TCTATCGCTTTCGAACCCCGGTGCACCTGTCGGAACGCAATGGGGAACACCGCTT 661  
 Db 601 TCTATCGCTTTCGAACCCCGGTGCACCTGTCGGAACGCAATGGGGAACACCGCTT 660  
 QY 662 TTTGGATGATTATGCAATGCTCTCCACATTTGATGCTTCCAGATTTCTGTTGGGAATACG 721  
 Db 661 TTTGGATGATTATGCAATGCTCTCCACATTTGATGCTTCCAGATTTCTGTTGGGAATACG 720  
 QY 722 CTGATAGCCTAACGTTTCATGATCAAAATTAATCTTAACCCCTACTTGACAGCAATA 781  
 Db 721 CTGATAGCCTAACGTTTCATGATCAAAATTAATCTTAACCCCTACTTGACAGCAATA 780  
 QY 782 TATTAACAGAGGAAGCTGCCCTGCTTAAACCTTTTATATCATCATTTAGCTTA 841  
 Db 781 TATTAACAGAGGAAGCTGCCCTGCTTAAACCTTTTATATCATCATTTAGCTTA 840  
 QY 842 CTTTTCATTAATTCGACGTTTCCAAATGACAACTTTTCATTTAAAGCTTTTAAACGAC 901  
 Db 841 CTTTTCATTAATTCGACGTTTCCAAATGACAACTTTTTCATTTTAAACGACTTTTAAACGAC 900  
 QY 902 AACTTGAGAAGATCAAAAACAACTAATTTTCGAGAGTCCAAAAGATGAGATTTCCCT 961  
 Db 901 AACTTGAGAAGATCAAAAACAACTAATTTTCGAGAGTCCAAAAGATGAGATTTCCCT 960  
 QY 962 CAATTTTACTGACGTTTATTCGACGATCTCTCCGATTTAGCTGCTCCAGTCAACACTA 1021  
 Db 961 CAATTTTACTGACGTTTATTCGACGATCTCTCCGATTTAGCTGCTCCAGTCAACACTA 1020  
 QY 1022 CAACAGAGATGAACGCAAAATTCGCGCTGAAGCTGATCTGCTTACTCAGATTAG 1081  
 Db 1021 CAACAGAGATGAACGCAAAATTCGCGCTGAAGCTGATCTGCTTACTCAGATTAG 1080  
 QY 1082 AAGGGATTTTCGATGTTGCTGTTTTCGCTTTTCCAAACAGCAAAATAACGGGTATTGT 1141  
 Db 1081 AAGGGATTTTCGATGTTGCTGTTTTCGCTTTTCCAAACAGCAAAATAACGGGTATTGT 1140  
 QY 1142 TTATAAATACTACTATTTCGACGATTTCTGCTAAAGAGAGGGGTATCTCTCGAGAAA 1201  
 Db 1141 TTATAAATACTACTATTTCGACGATTTCTGCTAAAGAGAGGGGTATCTCTCGAGAAA 1200  
 QY 1202 GAGGGCTGAAGCCCGG----- 1219  
 Db 1201 GATCGGACCGGCGAGGGGTTGCGGAAGAGGAGGCAACCCAAAAGCTGACCCCTTAG 1260  
 QY 1220 -----CCGAGGCCAGGCCCGAGGCCAG 1242  
 Db 1261 CCTACAGCAGTTTATCCCAATGTTGCGCGAGAGAGACCTTAGGCGCCAGGAGGTATG 1320  
 QY 1243 GAAGAAACAT----- 1252  
 Db 1321 AAGGGAAGATCTCCAGAAACTCCGAGCGATTTAAGGAATCAACCCCAATTAACCCCG 1380  
 QY 1253 -----GTATCATGGCCCCCGGAGAGGATAAATTTGTGCTTCCCCGGTGTG 1299  
 Db 1381 ACATCATATTAAAGGATGAAGAAACACCGGAGGAGGCTGATGACTCAGAGGTGTA 1440  
 QY 1300 A----- 1300  
 Db 1441 AGGCAAGTTGAACGCTTTGGCCATCTCGGTGATGAACACAGTGGCCAGGAGTGAACCTG 1500  
 QY 1301 ----- 1300  
 Db 1501 GGGTACCGAGGGCTGGAGCAAGATGCCCACTCAGAGGAGTCTCTGCACCTACGAGG 1560

QY 1301 ----- 1300  
Db 1561 GCGCGCAGTGGACATACACCGTCTGACCGGACCGCAGCAAGTACGGCATGCTGGCCC 1620  
QY 1301 ----- 1300  
Db 1621 GCCTGGCGGTGGAGCGCGCTTCGACTGGGTGTACTAGAGTCCAAAGGCACATACCCT 1680  
QY 1301 ----- 1300  
Db 1681 GCTCGGTGAAGCAGAGAACTCGGTGGCGGCAAAATCGGAGGCGTGGAGTGGCCAGG 1740  
QY 1301 ----- 1300  
Db 1741 ATTGTGGTTGAAGCCTTGCATATGTACAGTCCCAGAAAGTATCATCTGTCTCATCTCC 1800  
QY 1301 ----- 1300  
Db 1801 CCCAAAGCCCAAGGATGTGCTCACCAATTACTCTGACTCTTAAGTCAAGTGTGTGTGG 1860  
QY 1301 ----- 1300  
Db 1861 TAGACATCAGCAAGGATGATCCGAGGTCCAGTTCAGTGTGTTGTAGATGATGTGGAG 1920  
QY 1301 ----- 1300  
Db 1921 TGCACACAGCTCAGACGCAACACGGGAGAGAGAGTTCCAAAGCACTTTCGCTCAGTCA 1980  
QY 1301 ----- 1300  
Db 1981 GTGAATTTCCCATCATGACCAGGAGTGGCTCAATGGCAAGGAGTTCAAAATGCAAGGTCA 2040  
QY 1301 ----- 1300  
Db 2041 ACAGTGCAGCTTCCTGCCCCCATCGAGAAAACCATCTCCAAACCAAGGACAGCGCA 2100  
QY 1301 ----- CCG 1303  
Db 2101 AGGCTCCAGGTGPACCAATTCCACCTCCCAAGGAGCAGATGGCCAAAGGATAAAGTCA 2160  
QY 1304 CCCAGCAGTGCAGGAGAGAGTGTCTGTTTGTATGACAGTGTCCGGGATTCCTCGGT 1363  
Db 2161 GTCTGACTGTGATGATAACAGCTTCTCCCTGAAGACATTAAGTGTGAGTGGCAGTGA 2220  
QY 1364 GCTCCACCCCATGCGCCATCGAGAACA----- 1390  
Db 2221 ATGGCAGCCAGCGGAGAACTACAAGAACACTCAGCCCATCATGGACACAGATGGCTTT 2280  
QY 1391 ----- 1390  
Db 2281 ACTTCGTCTACAGCAAGCTCAATGTGCAGAGAGCAACTGGGAGGCGAGAAATCTTTCA 2340  
QY 1391 ----- CTCAAGAGAGAGATGTCCCTTTCTAACT 1418  
Db 2341 CCGTCTGTGTATCAGGGCCCTGCACACCACTACTGAGAAGAGCTCTCCCACT 2400  
QY 1419 AGTGGCGTAGA-----ATTCCCTAGGCGCGCGGAATTA 1453  
Db 2401 CTCCTGTAATATGATCCCAAGTGTCTTGGAGCCCTCTGTGCTCTACAGCGCGCGGAATTA 2460  
QY 1454 ATTCGCTTTAGACATGACTTCTCAGTTTCAGTTGGGCACTTACAGAGAGACCGGTCT 1513  
Db 2461 ATTCGCTTTAGACATGACTTCTCAGTTTCAGTTGGGCACTTACAGAGAGACCGGTCT 2520  
QY 1514 TGCTAGATCTTAATCAAGAGAGATGTGAGATGCCATTTGCTGAGAGATGAGAGCTTCAT 1573  
Db 2521 TGCTAGATCTTAATCAAGAGAGATGTGAGATGCCATTTGCTGAGAGATGAGAGCTTCAT 2580  
QY 1574 TTTTGATACCTTTTATTTGTAACCTATATAGTATAGATTTTGTGATTTTGTCTTTTTC 1633  
Db 2581 TTTTGATACCTTTTATTTGTAACCTATATAGTATAGATTTTGTGATTTTGTCTTTTTC 2640

QY 1634 TTCTCGTACGAGCTTGTCTGATCAGCTATCTCGAGCTGATGAATATCTTGTGGTAG 1693  
Db 2641 TTCTCGTACGAGCTTGTCTGATCAGCTATCTCGAGCTGATGAATATCTTGTGGTAG 2700  
QY 1694 GGGTTTGGAAATCATTGAGTTTGTATTTTCTTGGTATTTCCCACTCCTCTTTCAGA 1753  
Db 2701 GGGTTTGGAAATCATTGAGTTTGTATTTTCTTGGTATTTCCCACTCCTCTTTCAGA 2760  
QY 1754 GTACAGAGATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATGCGGT 1813  
Db 2761 GTACAGAGATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATGCGGT 2820  
QY 1814 AGTTTATCACAGTTAAATTTGCTAAAGCAGTCCAGCACCGTGTATGAAATCTAAACAATCG 1873  
Db 2821 AGTTTATCACAGTTAAATTTGCTAAAGCAGTCCAGCACCGTGTATGAAATCTAAACAATCG 2880  
QY 1874 CTCATCGTCTCCTCGGCACCGTCACTCGTGTGATGCTGTAGGATAGGTTGGTTATGCGG 1933  
Db 2881 CTCATCGTCTCCTCGGCACCGTCACTCGTGTGATGCTGTAGGATAGGTTGGTTATGCGG 2940  
QY 1934 GTACTGCCGGGCTCTTGGCGGATATCGTCCATTCGACAGCATCGCCAGTCACTATGCG 1993  
Db 2941 GTACTGCCGGGCTCTTGGCGGATATCGTCCATTCGACAGCATCGCCAGTCACTATGCG 3000  
QY 1994 GTGCTGTAGCCTATATGCGTTGATGCAATTTCTATGCGCACCGCTTCTCGGAGCATG 2053  
Db 3001 GTGCTGTAGCCTATATGCGTTGATGCAATTTCTATGCGCACCGCTTCTCGGAGCATG 3060  
QY 2054 TCCGACCGTTTGGCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2113  
Db 3061 TCCGACCGTTTGGCGCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
QY 2114 TAGCGATCANGCGGACCAACCCCGTCTGTGATGCTATCGAATCTAAATGTAGTTAAA 2173  
Db 3121 TAGCGATCANGCGGACCAACCCCGTCTGTGATGCTATCGAATCTAAATGTAGTTAAA 3180  
QY 2174 ATCTCTAATAATTAATAAGTCCAGTTTCTCCATACGAACTTAAACAGATTCGCGTG 2233  
Db 3181 ATCTCTAATAATTAATAAGTCCAGTTTCTCCATACGAACTTAAACAGATTCGCGTG 3240  
QY 2234 AGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAAATATGTTCACTCC 2293  
Db 3241 AGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAAATATGTTCACTCC 3300  
QY 2294 CTCAGAGTTAGCTTGTGAGATGATGAATTTCTGGAAGTTCAGTGTAACTCCGCT 2353  
Db 3301 CTCAGAGTTAGCTTGTGAGATGATGAATTTCTGGAAGTTCAGTGTAACTCCGCT 3360  
QY 2354 GTATTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGTTACCGGAGGAGTAACTTCC 2413  
Db 3361 GTATTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGTTACCGGAGGAGTAACTTCC 3420  
QY 2414 ACACTCTCTGAGAGTAGGCAACCAACACACAGATCCAGGTTGTTGTTGTTGTTGTTGTTG 2473  
Db 3421 ACACTCTCTGAGAGTAGGCAACCAACACACAGATCCAGGTTGTTGTTGTTGTTGTTGTTG 3480  
QY 2474 ATAAGAAGAGCATTTCTGATTTGAGGATCAAGTGTTCAGGAGGCTACTGATGGACAT 2533  
Db 3481 ATAAGAAGAGCATTTCTGATTTGAGGATCAAGTGTTCAGGAGGCTACTGATGGACAT 3540  
QY 2534 TTCCAAAGCCTCTGATAGTTGCAACCGATAGGTTGTAGAGTGTGCAATFACACTTCCG 2593  
Db 3541 TTCCAAAGCCTCTGATAGTTGCAACCGATAGGTTGTAGAGTGTGCAATFACACTTCCG 3600  
QY 2594 TACAATTTCAACCTTGGCACTGCACAGCTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2653  
Db 3601 TACAATTTCAACCTTGGCACTGCACAGCTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3660  
QY 2654 AAGCTCCTTGTCTCATATGACAGGCAACAGAAATCACTCGGGAATCAATACCATGTTTC 2713  
Db 3661 AAGCTCCTTGTCTCATATGACAGGCAACAGAAATCACTCGGGAATCAATACCATGTTTC 3720  
QY 2714 ASCTTGACAGAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTATTTTATCAGCAATAAC 2773

Db 3721 AGCTTGAG-CAGAGGCTGAGGCAACGAATCTGGATCAGGCTATTTATCAGCAATAAC 3779  
Qy 2774 TAGAACTTCAGAGGCCAGCAGGATGTCAATCTACACAGGCGTATGTGTCATTTTG 2833  
Db 3780 TAGAACTTCAGAGGCCAGCAGGATGTCAATCTACACAGGCGTATGTGTCATTTTG 3839  
Qy 2834 AACCATCATCTTGGCAGCAGTAACGAACCTGTTCTCTGGACCAATATTTTGTACACTT 2893  
Db 3840 AACCATCATCTTGGCAGCAGTAACGAACCTGTTCTCTGGACCAATATTTTGTACACTT 3899  
Qy 2894 AGGAACAGTCTTCTGTTCCGTAAGCCATAGCAGTACTGCTGGGGCGCTCTGCTAGCAC 2953  
Db 3900 AGGAACAGTCTTCTGTTCCGTAAGCCATAGCAGTACTGCTGGGGCGCTCTGCTAGCAC 3959  
Qy 2954 GATACACTTAGCACCAACCTTGTGGCAACCTAGTACTCTCTGGGTAAAGGTACCATC 3013  
Db 3960 GATACACTTAGCACCAACCTTGTGGCAACCTAGTACTCTCTGGGTAAAGGTACCATC 4019  
Qy 3014 CTTCCTAGGTGAGATGCAAAAACAAATTTCTTTGCAACAGCAACTTTTGGCAGGAACACC 3073  
Db 4020 CTTCCTAGGTGAGATGCAAAAACAAATTTCTTTGCAACAGCAACTTTTGGCAGGAACACC 4079  
Qy 3074 CAGCATAGGGAAGTGGGAAGCAGAAATTCGGTTCACACAGGAATATAGAGGCCAATTT 3133  
Db 4080 CAGCATAGGGAAGTGGGAAGCAGAAATTCGGTTCACACAGGAATATAGAGGCCAATTT 4139  
Qy 3134 CTCAATAGTCTTTCGAAACAGAGACGACTACACAGGCAAGTCTCAACTTGCACAGT 3193  
Db 4140 CTCAATAGTCTTTCGAAACAGAGACGACTACACAGGCAAGTCTCAACTTGCACAGT 4199  
Qy 3194 CTCCGTTAGTTGAGCTTCAGAAATTCCTGACGTTATCTATAGAGATCAATGGCTCT 3253  
Db 4200 CTCCGTTAGTTGAGCTTCAGAAATTCCTGACGTTATCTATAGAGATCAATGGCTCT 4259  
Qy 3254 CTTCACGTTATCTGCAATGTCATAGTTCTCTCTGGAAAGAGCTTCTAAACAGAGTGT 3313  
Db 4260 CTTCACGTTATCTGCAATGTCATAGTTCTCTCTGGAAAGAGCTTCTTAACAGAGTGT 4319  
Qy 3314 CTTCGAAAGCACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTTTGTCAACATTTG 3373  
Db 4320 CTTCGAAAGCACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTTTGTCAACATTTG 4379  
Qy 3374 AGCAATTTGTCGAATTTGTTTGTGACATTAATTCATATCTGTTCCGTTCTTGATAGG 3433  
Db 4380 AGCAATTTGTCGAATTTGTTTGTGACATTAATTCATATCTGTTCCGTTCTTGATAGG 4439  
Qy 3434 AGCAATTTGTCGAATTTGTTTGTGACATTAATTCATATCTGTTCCGTTCTTGATAGG 3493  
Db 4440 AGCAATTTGTCGAATTTGTTTGTGACATTAATTCATATCTGTTCCGTTCTTGATAGG 4499  
Qy 3494 TTCAATAGCACTTCAGAGGCACTCTTTAGGTTTGGATTCCTTTAGGTTTGTTCCTT 3553  
Db 4500 TTCAATAGCACTTCAGAGGCACTCTTTAGGTTTGGATTCCTTTAGGTTTGTTCCTT 4559  
Qy 3554 GGTGATCTCTGGCTTGGCATCTCTTCTCTAGTGACCTTTAGGACTTCATATCCAG 3613  
Db 4560 GGTGATCTCTGGCTTGGCATCTCTTCTCTAGTGACCTTTAGGACTTCATATCCAG 4619  
Qy 3614 GTTCTCTCCACCTCGTCCAAAGTCACACCGTACTTTGGCACTTAATGCAAAATA 3673  
Db 4620 GTTCTCTCCACCTCGTCCAAAGTCACACCGTACTTTGGCACTTAATGCAAAATA 4679  
Qy 3674 AAATAAGTCAGACATTCAGAGCTATCTTCTTGGATTTAGCTTCTGCAAGTTTATC 3733  
Db 4680 AAATAAGTCAGACATTCAGAGCTATCTTCTTGGATTTAGCTTCTGCAAGTTTATC 4739  
Qy 3734 AGCTTCTCTCCCTAATTTAGCGTTC-AACAAAACCTTCTGCTGCTCAATAACCGTTTGGTAT 3792  
Db 4740 AGCTTCTCTCCCTAATTTAGCGTTCACACAAACTTCTGCTCAATAACCGTTTGGTAT 4799  
Qy 3793 AAGAACTTCTGAGGATTCCTCTTAGATCCCAAGGTTGGTTCCTATGGCTCTAAGAC 3852

Db 4800 AAGAACTTCTGAGGATTCCTCTTACGATCCCAAGGT-CCTTCCATGGCTCTAAGAC 4858  
Qy 3853 CTTTGTATTTGGCAAAAACAGGAAGTGCCTTCCAAAGTGACAGAAACCAACACCTGTTGT 3912  
Db 4859 CTTTGTATTTGGCAAAAACAGGAAGTGCCTTCCAAAGTGACAGAAACCAACACCTGTTGT 4918  
Qy 3913 CAAACCAAAATTTCAAGCGATCTCATCAATCCAAATTCGATACCCAGCAACTTTGAG 3972  
Db 4919 CAAACCAAAATTTCAAGCGATCTCCATCAATCCAAATTCGATACCCAGCAACTTTGAG 4978  
Qy 3973 TTGCTCCAGATGATAGCACCTTTATACCAAAACCGTGACGAGATTTGGTAGACTCCAG 4032  
Db 4979 TTGCTCCAGATGATAGCACCTTTATACCAAAACCGTGACGAGATTTGGTAGACTCCAG 5038  
Qy 4033 TTTGTCTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCGCCCAACGAGT 4092  
Db 5039 TTTGTCTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCGCCCAACGAGT 5098  
Qy 4093 AATTGAAAGTGCAGCACCAAGTAGTCAATAGACCATCTGGGGCGGTCACTAGTCAAAAG 4152  
Db 5099 AATTGAAAGTGCAGCACCAAGTAGTCAATAGACCATCTGGGGCGGTCACTAGTCAAAAG 5158  
Qy 4153 ACGCCCAACAAAATTTCACTGACAGGGAACCTTTTGTGACATCTTCCAGAAAGTTCGTATTCAG 4212  
Db 5159 ACGCCCAACAAAATTTCACTGACAGGGAACCTTTTGTGACATCTTCCAGAAAGTTCGTATTCAG 5218  
Qy 4213 TAGTCAATTTGCCGAGCATCAATTAATGGGATTTATACAGAGCAACAGTGGAGTCAACAT 4272  
Db 5219 TAGTCAATTTGCCGAGCATCAATTAATGGGATTTATACAGAGCAACAGTGGAGTCAACAT 5278  
Qy 4273 CTACCAACTTTGGCGTCTCAGAAAAAGCATAAACAGTCTTACTACCGCCATTTAGTGAAC 4332  
Db 5279 CTACCAACTTTGGCGTCTCAGAAAAAGCATAAACAGTCTTACTACCGCCATTTAGTGAAC 5338  
Qy 4333 TTTTCAAAATCGCCAGTGGAGAAAGGCAAGGATCTAGCATTTAGCGGGCAAGG 4392  
Db 5339 TTTTCAAAATCGCCAGTGGAGAAAGGCAAGGATCTAGCATTTAGCGGGCAAGG 5398  
Qy 4393 ATGCAACTTTTATCAACAGGCTCTATAGATAACCTTAGCCCTGGGATCATCTTTTGA 4452  
Db 5399 ATGCAACTTTTATCAACAGGCTCTATAGATAACCTTAGCCCTGGGATCATCTTTTGA 5458  
Qy 4453 CAACCTTTCTGCAAAATCTAGGTCGAAATCTCTTATTTGATACCAATTTAGTCAACT 4512  
Db 5459 CAACCTTTCTGCAAAATCTAGGTCGAAATCTCTTATTTGATACCAATTTAGTCAACT 5518  
Qy 4513 TGACCAAGTTTGTGATCAGCTCTCTCAAAATTTGCTCTGTAACGAGTACTCAACTTGA 4572  
Db 5519 TGACCAAGTTTGTGATCAGCTCTCTCAAAATTTGCTCTGTAACGAGTACTCAACTTGA 5578  
Qy 4573 CATTAACTTTGAAGCTCAGTCTGATTTGATGAACCTTTGATCAGGTTGTGAGTGTGAGCAG 4632  
Db 5579 CATTAACTTTGAAGCTCAGTCTGATTTGATGAACCTTTGATCAGGTTGTGAGTGTGAGCAG 5638  
Qy 4633 CATAGGGAACACAGGCTTTTCTTACCAACTCAAGGAATTTCAAACTCTGCAACACTTG 4692  
Db 5639 CATAGGGAACACAGGCTTTTCTTACCAACTCAAGGAATTTCAAACTCTGCAACACTTG 5698  
Qy 4693 CGTATGAGGTAGCAAGGAAATGTCTATCTTGAAGTCGACAGTACTGAGTGTGAGTGTGAG 4752  
Db 5699 CGTATGAGGTAGCAAGGAAATGTCTATCTTGAAGTCGACAGTACTGAGTGTGAGTGTGAG 4812  
Qy 4753 AAATTTCTGAAGCCGATTTTATTTATCAGTCACTGATCTAGGAGATCTCTACGCGG 4872  
Db 5759 AAATTTCTGAAGCCGATTTTATTTATTTATCAGTCACTGATCTAGGAGATCTCTACGCGG 5818  
Qy 4813 GACGATCTGGGCGACCTGCGGTCGATCACCGGCGCCACAGGTGCGGTTCTGCGG 4872  
Db 5819 GACGATCTGGGCGACCTGCGGTCGATCACCGGCGCCACAGGTGCGGTTCTGCGG 5878  
Qy 4873 CCTATATCGCGACATCACCGGATGGGAAGATCGGGTCTGCGGCTCGGCTCATGAGCG 4932  
Db 5879 CCTATATCGCGACATCACCGGATGGGAAGATCGGGTCTGCGGCTCGGCTCATGAGCG 5938



Db 8099 CGCGTTTCGGTGATGACGGTGAAGAACCTCTGACACATGACGCTCCCGAGACGGTCAAG 8158  
Qy 5913 CTTGCTGTAAAGCGGATCCGGGAGCAGACAAGCCCGTCAGGGCGGTCAGCGGGTTTG 5972  
Db 8159 CTTGCTGTAAAGCGATCCGGGAGCAGACAAGCCCGTCAGGGCGGTCAGCGGGTTTG 8218  
Qy 5973 CGGGTGTCCGGGCGCAGCATGACCCAGTCACCGTAGCGATAGCGAGTGTATCTGGCT 6032  
Db 8219 CGGGTGTCCGGGCGCAGCATGACCCAGTCACCGTAGCGATAGCGAGTGTATCTGGCT 8278  
Qy 6033 TAACTATCGGCATCAGAGCAGATGTACTGAGAGTGCACCATATGCGGTTGGAATACC 6092  
Db 8279 TAACTATCGGCATCAGAGCAGATGTACTGAGAGTGCACCATATGCGGTTGGAATACC 8338  
Qy 6093 GCACAGATGCTGAGGAGAAATACCGCATCAGGGCTCTTCGCTTCCTCGCTCACTGA 6152  
Db 8339 GCACAGATGCTGAGGAGAAATACCGCATCAGGGCTCTTCGCTTCCTCGCTCACTGA 8398  
Qy 6153 CTCGCTCGCTCGCTCGCTCGCTCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAT 6212  
Db 8399 CTCGCTCGCTCGCTCGCTCGCTCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAT 8458  
Qy 6213 ACGGTTATCCACAGATCAGGGATTAACGACAGGAAGAACATGTAGCAAAAGCGCAGCA 6272  
Db 8459 ACGGTTATCCACAGATCAGGGATTAACGACAGGAAGAACATGTAGCAAAAGCGCAGCA 8518  
Qy 6273 AAAGCCAGGACCGTAAGAGCGCGCTGCTGCGGTTTTCATAGGCTCCGCCCCC 6332  
Db 8519 AAAGCCAGGACCGTAAGAGCGCGCTGCTGCGGTTTTCATAGGCTCCGCCCCC 8578  
Qy 6333 TCACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGGGAACCCGACAGCACTATA 6392  
Db 8579 TCACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGGGAACCCGACAGCACTATA 8638  
Qy 6393 AAGATACAGGGGTTTCCCTCGGAGGCTCCCTCGTGGCTCTCTGTCGACCTGCC 6452  
Db 8639 AAGATACAGGGGTTTCCCTCGGAGGCTCCCTCGTGGCTCTCTGTCGACCTGCC 8698  
Qy 6453 GCTTACCGGATACCTGTGCGCTTCTCCCTCGGAGGCTGGGCTTCTCATAGTCT 6512  
Db 8699 GCTTACCGGATACCTGTGCGCTTCTCCCTCGGAGGCTGGGCTTCTCATAGTCT 8758  
Qy 6513 ACGCTGTAGGTATCAGTTCGGGTGAGTTCGCTCAAGTGGGCTGTCACGA 6572  
Db 8759 ACGCTGTAGGTATCAGTTCGGGTGAGTTCGCTCAAGTGGGCTGTCACGA 8818  
Qy 6573 ACCCCCGTTACCCCGGCTGCGCTTATCCGTAACCTATCGTTCGATCCAAACC 6632  
Db 8819 ACCCCCGTTACCCCGGCTGCGCTTATCCGTAACCTATCGTTCGATCCAAACC 8878  
Qy 6633 GGTAAACACGACTTATCGCCACTGSCAGCAGCCACTGGTAACAGGATTAGCAGCGAG 6692  
Db 8879 GGTAAACACGACTTATCGCCACTGSCAGCAGCCACTGGTAACAGGATTAGCAGCGAG 8938  
Qy 6693 GTATGTAGGGGTGCTACAGAGTCTTGAAGTGGTGGCTTAAGTGGTGGTGGTGGT 6752  
Db 8939 GTATGTAGGGGTGCTACAGAGTCTTGAAGTGGTGGCTTAAGTGGTGGTGGTGGT 8998  
Qy 6753 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTCCTTCGGAAGAGGTTGGTAG 6812  
Db 8999 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTCCTTCGGAAGAGGTTGGTAG 9058  
Qy 6813 CTTTGTATCGGCAAAACAAACCCAGCTGCTAGCGGTGGTGGTGGTGGTGGTGGTGGT 6872  
Db 9059 CTTTGTATCGGCAAAACAAACCCAGCTGCTAGCGGTGGTGGTGGTGGTGGTGGTGGT 9118  
Qy 6873 GATTACCGCAGAAAAAGGATCTCAGAGATCCTTTGATCTTTTCTACGGGTCTGA 6932  
Db 9119 GATTACCGCAGAAAAAGGATCTCAGAGATCCTTTGATCTTTTCTACGGGTCTGA 9178  
Qy 6933 CGCTCAGTGAACGAAACTCAGCTTAAGGGATTTTGGTCATGAGATTAATAAAGGAT 6992

Db 9179 CGCTCAGTGGAAACAACTCAGCTTAAGGGATTTTGGTCATGAGATTAATAAAGGAT 9238  
Qy 6993 CTTACCTAGATCCCTTTTAAATTAATAATGAAGTTTAAATCAATCAATGAAGTATATGA 7052  
Db 9239 CTTACCTAGATCCCTTTTAAATTAATAATGAAGTTTAAATCAATCAATGAAGTATATGA 9298  
Qy 7053 GTAAACTTGGTCTGACAGTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGATCTG 7112  
Db 9299 GTAAACTTGGTCTGACAGTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGATCTG 9358  
Qy 7113 TCTATTTGGTTCATCCATAGTTCCTGACTCCCGTCTGCTAGATTAATCAATCAATGAAG 7172  
Db 9359 TCTATTTGGTTCATCCATAGTTCCTGACTCCCGTCTGCTAGATTAATCAATCAATGAAG 9418  
Qy 7173 GGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGCAGAGCCCAAGCTCACCAGCTCC 7232  
Db 9419 GGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGCAGAGCCCAAGCTCACCAGCTCC 9478  
Qy 7233 AGATTTATCAGCAATAAACCAGCAGCGAGGCGCGAGCGAGGAGTGGTCTCTCAAC 7292  
Db 9479 AGATTTATCAGCAATAAACCAGCAGCGAGGCGCGAGCGAGGAGTGGTCTCTCAAC 9538  
Qy 7293 TTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTTAGAGTAAGTAGTTCGCC 7352  
Db 9539 TTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTTAGAGTAAGTAGTTCGCC 9598  
Qy 7353 AGTTAATAGTTTCGCAACGTTGTCGCAATGCTGAGGCACTGCTGGTTCAGCTCGTC 7412  
Db 9599 AGTTAATAGTTTCGCAACGTTGTCGCAATGCTGAGGCACTGCTGGTTCAGCTCGTC 9658  
Qy 7413 GTTTGTATGGCTTCATTCAGCTCCCGTCCCAAGCATCAAGCGAGTTCATGATCCCC 7472  
Db 9659 GTTTGTATGGCTTCATTCAGCTCCCGTCCCAAGCATCAAGCGAGTTCATGATCCCC 9718  
Qy 7473 CATGTTGTGCAAAAAAGCGGTAGCTCTTCGCTCCCTCCGATCGTTGTTCAGAAAGTAA 7532  
Db 9719 CATGTTGTGCAAAAAAGCGGTAGCTCTTCGCTCCCTCCGATCGTTGTTCAGAAAGTAA 9778  
Qy 7533 GGGCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCCTTACTGTCTAGTCC 7592  
Db 9779 GGGCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCCTTACTGTCTAGTCC 9838  
Qy 7593 ATCCGTAAGATGCTTTTCTGCTGATGCTGCTCAACCAAGTCAATCTTGAGAATAGTG 7652  
Db 9839 ATCCGTAAGATGCTTTTCTGCTGATGCTGCTCAACCAAGTCAATCTTGAGAATAGTG 9898  
Qy 7653 TATCGCGCAGCGAGTTGCTCTTCGCGCGGTCAACACGGGATAATACCGCGCACATAG 7712  
Db 9899 TATCGCGCAGCGAGTTGCTCTTCGCGCGGTCAACACGGGATAATACCGCGCACATAG 9958  
Qy 7713 CAGAACTTTAAAGTGTCTCATTTGGAACACGTTCTTCGGGGGCAAACTCTCAAGCAT 7772  
Db 9959 CAGAACTTTAAAGTGTCTCATTTGGAACACGTTCTTCGGGGGCAAACTCTCAAGCAT 10018  
Qy 7773 CTTACCCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCACTTCCTCAGC 7832  
Db 10019 CTTACCCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCACTTCCTCAGC 10078  
Qy 7833 ATCTTTTACTTTACCGAGCTTCGCTGAGCAAAACAGGAGGCAAAATCGCGCAAA 7892  
Db 10079 ATCTTTTACTTTACCGAGCTTCGCTGAGCAAAACAGGAGGCAAAATCGCGCAAA 10138  
Qy 7893 AAAGGGAATAGGGGACACGAAATGTTGAATACTCATACTCTTCTCTTTTCAATATTA 7952  
Db 10139 AAAGGGAATAGGGGACACGAAATGTTGAATACTCATACTCTTCTCTTTTCAATATTA 10198  
Qy 7953 TTGAAGCATTTATCAGGGTATTTGCTCATGAGCGGATACATATTTGAAATGTTATAGAA 8012  
Db 10199 TTGAAGCATTTATCAGGGTATTTGCTCATGAGCGGATACATATTTGAAATGTTATAGAA 10258  
Qy 8013 AAATAACCAATAGGGTTCGCGCAGATTTCCCGGAAAGTGCACCTGACGCTAAGA 8072  
Db 10259 AAATAACCAATAGGGTTCGCGCAGATTTCCCGGAAAGTGCACCTGACGCTAAGA 10318

QY 8073 AACCAATTATCATGACATTAACTATAAAATAGGGGTATCAGAGGCCCTTTCGTCT 8132  
 Db 10319 AACCAATTATCATGACATTAACTATAAAATAGGGGTATCAGAGGCCCTTTCGTCT 10378  
 QY 8133 TCAAGAAATTAATTCATGTTTGACAGCTTATCATCGATAAGCTGACATCATGTTGGTATT 8192  
 Db 10379 TCAGAAATTAATTCATGTTTGACAGCTTATCATCGATAAGCTGACATCATGTTGGTATT 10438  
 QY 8193 GTGAATAGACGACATCGGGAACACTGAAATAAACAAGTTATTATTTCG 8241  
 Db 10439 GTGAATAGACGACATCGGGAACACTGAAATAAACAAGTTATTATTTCG 10487

RESULT 15  
 AX417107  
 LOCUS AX417107 10491 bp DNA linear PAT 14-JUN-2002  
 DEFINITION Sequence 32 from Patent WO0198344.  
 ACCESSION AX417107  
 VERSION AX417107.1 GI:21445694  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 unclassified.

REFERENCE 1  
 AUTHORS Ling, L. E. and Sanicola-Nadel, M.  
 TITLE Angiogenesis-modulating compositions and uses  
 JOURNAL Patent: WO 0198344-A 32 27-DEC-2001;  
 BIOGEN, INC. (US)  
 FEATURES  
 Location/Qualifiers  
 1..10491  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"  
 /note="Plasmid pUB114"

ORIGIN

Query Match 68.8%; Score 5665.8; DB 6; Length 10491;  
 Best Local Similarity 77.2%; Pred. No. 0;  
 Matches 8100; Conservative 0; Mismatches 137; Indels 2252; Gaps 10;

QY 2 GATCTAAATCCAAAGACGAAGCTTGAATGAACCTTTTGGCATCCGACATCCACAGG 61  
 Db 1 GATCTAAATCCAAAGACGAAGCTTGAATGAACCTTTTGGCATCCGACATCCACAGG 60  
 QY 62 TCATTCTCACACATAAGTGCACAAACGCAACGAGGAGGATACACTAGCAGCAGACCGTT 121  
 Db 61 TCATTCTCACACATAAGTGCACAAACGCAACGAGGAGGATACACTAGCAGCAGACCGTT 120  
 QY 122 GCAACGAGACCTCCACTCTCTCTCTCTCAACACCCACTTTGGCATCGAATAACCA 181  
 Db 121 GCAACGAGACCTCCACTCTCTCTCTCAACACCCACTTTGGCATCGAATAACCA 180  
 QY 182 GCCAGTTATTGGGCTTCATTTGGAGCTCGCTCATTTCCAAATTCCTTATTAGGCTACTAA 241  
 Db 181 GCCAGTTATTGGGCTTCATTTGGAGCTCGCTCATTTCCAAATTCCTTATTAGGCTACTAA 240  
 QY 242 CACCATGACTTATTAGCTGTCTATCTGCTGCCCCCTCGGCGAGGTTCAATGTTGTTTAT 301  
 Db 241 CACCATGACTTATTAGCTGTCTATCTGCTGCCCCCTCGGCGAGGTTCAATGTTGTTTAT 300  
 QY 302 TTCCGAATGCAACAGCTCCGCAATACACCGCAACATCACTCCAGATGAGGCTTCTGA 361  
 Db 301 TTCCGAATGCAACAGCTCCGCAATACACCGCAACATCACTCCAGATGAGGCTTCTGA 360  
 QY 362 GTGTGGGTCAAATAGTTTCATGTTCCCAATGCGCCAAATCGACAGTTTAAACGCTG 421  
 Db 361 GTGTGGGTCAAATAGTTTCATGTTCCCAATGCGCCAAATCGACAGTTTAAACGCTG 420  
 QY 422 TCTTGGAACTTAATGACAAAGCGGTGATCTCATCCAGATGAACCTAAGTTGGTTCGT 481  
 Db 421 TCTTGGAACTTAATGACAAAGCGGTGATCTCATCCAAAGATGAACCTAAGTTGGTTCGT 480

QY 482 TGAATGCTAAACGGCCAGTTGGTCAAAAGAAACCTTCCAAAGTCCCATACCGTTTCTC 541  
 Db 481 TGAATGCTAAACGGCCAGTTGGTCAAAAGAAACCTTCCAAAGTCCCATACCGTTTCTC 540  
 QY 542 TTGTTTGGTATTGATGACGAATGCTCAAAAATAATCTCAATTAATGCTTAGCGAGTCTC 601  
 Db 541 TTGTTTGGTATTGATGACGAATGCTCAAAAATAATCTCAATTAATGCTTAGCGAGTCTC 600  
 QY 602 TCTATCGCTTCTGAACCCCGGTGCACTGCTGCCGAAACGCAAAATGGGAAACACCCGTT 661  
 Db 601 TCTATCGCTTCTGAACCCCGGTGCACTGCTGCCGAAACGCAAAATGGGAAACACCCGTT 660  
 QY 662 TTGATGATGATTATGCAATGCTCCACATGCTATGCTTCAAGATTCCTGGTGGGAACTCTG 721  
 Db 661 TTGATGATGATTATGCAATGCTCCACATGCTATGCTTCAAGATTCCTGGTGGGAACTCTG 720  
 QY 722 CTGATAGCCTAAACGTTTCATGATCAAAATTAACCTTCTAACCCCTACTTGCAGCAATA 781  
 Db 721 CTGATAGCCTAAACGTTTCATGATCAAAATTAACCTTCTAACCCCTACTTGCAGCAATA 780  
 QY 782 TATAACAGAGGAAGCTGCCCTGCTTAAACCTTTTATCATCATTAATAGCTTA 841  
 Db 781 TATAACAGAGGAAGCTGCCCTGCTTAAACCTTTTATCATCATTAATAGCTTA 839  
 QY 842 CTTTCATAATTCGAGTGGTTCCAATTCAGACGTTTGTATTTAACACTTTTAAACGAC 901  
 Db 840 CTTTCATAATTCGAGTGGTTCCAATTCAGACGTTTGTATTTAACGACTTTTAAACGAC 899  
 QY 902 AACTTGAGAGATCAAAAACAACTAATTTTCGAGGATCCAAACGATGAGATTCCTT 961  
 Db 900 AACTTGAGAGATCAAAAACAACTAATTTTCGAGGATCCAAACGATGAGATTCCTT 959  
 QY 962 CAATTTTACTCGAGTTTATTTTCGAGCATCTCCGCAATAGCTCTCCAGTCAACACTA 1021  
 Db 960 CAATTTTACTCGAGTTTATTTTCGAGCATCTCCGCAATAGCTCTCCAGTCAACACTA 1019  
 QY 1022 CAACAGAGATGAAACGGCAAAATTCGGCTGAAGCTGTCAATCGGTTACTCAGATTAG 1081  
 Db 1020 CAACAGAGATGAAACGGCAAAATTCGGCTGAAGCTGTCAATCGGTTACTCAGATTAG 1079  
 QY 1082 AAGGGATTTCCATGTTGCTGTTTTCGCAATTTTCCAAACAGACAAATTAACGGTTATTGT 1141  
 Db 1080 AAGGGATTTCCATGTTGCTGTTTTCGCAATTTTCCAAACAGACAAATTAACGGTTATTGT 1139  
 QY 1142 TTATAAATACTACTATGCTCAGCATGCTGCTAAAGAAAGGGGTATCTCTCGAGAAAA 1201  
 Db 1140 TTATAAATACTACTATGCTCAGCATGCTGCTAAAGAAAGGGGTATCTCTCGAGAAAA 1199  
 QY 1202 GAGAGCTGAAGCCCGAGG----- 1219  
 Db 1200 GATGCGACCGGCGAGGGGTTTCGGGAAGAGAGAGGACCCCAAAAGCTGACCCCTTTAG 1259  
 QY 1220 -----CCAGGCCAGGCCAGGCCAG 1242  
 Db 1260 CCTACAGCAGTTTATCCCAATGTGGCGGAGAGACCTTAGCGCCAGCGAAGGTATG 1319  
 QY 1243 GAAGAAACAT----- 1252  
 Db 1320 AAGGGAAGATCTCCAGAACTCCGAGCGATTTAAGAACTCACCCCCAATTAACACCCG 1379  
 QY 1253 -----GTATCATGGCCCCCGGAGAGGATAAATTTGGCTTCCCCGGTCTC 1299  
 Db 1380 ACATCATATTTAAGGATGAAGAAACACCGGAGCGACAGGCTGATGACTCAGAGGTCTA 1439  
 QY 1300 A----- 1300  
 Db 1440 AGGACAAGTTGAACGTTTGGCCATCTCGGTGATGAACAGTGGCCAGGAGTGAACCTGC 1499  
 QY 1301 ----- 1300  
 Db 1500 GGTGACCGGGCTGGGACGAGATGCCACCCTCAGAGGATCTCTGCATACGAGG 1559  
 QY 1301 ----- 1300



Db 1560 GCCCGCAGTGGACATCACCAGTCTGACCGGACCGCAGCAAGTACGGCATGCTGGCCC 1619  
Qy 1301 ----- 1300  
Db 1620 GCGTGGCGGTGGAGCGCGCTTCGACTGGGTGTAACAGTCCAAAGCACAATATCCACT 1679  
Qy 1301 ----- 1300  
Db 1680 GCTCGGTGAACGACAGAGAACTCGGTGGCGGCCAAATCGGAGGCGTCAAGTGGCCAGGG 1739  
Qy 1301 ----- 1300  
Db 1740 ATTGGTGTGAAGCCTTGATATGTACAGTCCAGAGATATCTGTCTTCATCTCC 1799  
Qy 1301 ----- 1300  
Db 1800 CCCCAGCCAAAGGATGTGTCACCAATTAATCTGACTTAAAGTCAAGTCAAGTGTGTGG 1859  
Qy 1301 ----- 1300  
Db 1860 TAGACATCAGCAAGGATGATCCGAGGTCAGTTCAGCTGGTTGTAGATGATGTGGAGG 1919  
Qy 1301 ----- 1300  
Db 1920 TGCACACAGCTCAGACGCAACCAACGGAAGAGCAGTTCCAAAGCACTTTCCGCTCAGTCA 1979  
Qy 1301 ----- 1300  
Db 1980 GTGAATCTCCATCATGCAACGAGGACTGGCTCAATGGCAAGAGTTCAAATGCAGGGTCA 2039  
Qy 1301 ----- 1300  
Db 2040 ACAGTGCAGCTTTCCTGCGCCCATCGAGAAACCATCTCCAAACCAAGGACAGACCGA 2099  
Qy 1301 -----CGG 1303  
Db 2100 AGGCTCCACAGGTGTACACATTCACCTCCCAAGAGCAGATGGCCAGAGTAAGTCA 2159  
Qy 1304 CCCAGCAGTGCAGGAGAGAGTGTGTTTGTATGACAGTGTCCGGGGATTTCCCGTGGT 1363  
Db 2160 GTCTGACCTGCATGATACAGACTTCTCCCTGAAGACATTACTGTGGAGTGGCAGTGG 2219  
Qy 1364 GCTTCCACCCATGGCCATCGAGACA ----- 1390  
Db 2220 ATGGGCAGCGCAGGAGAACTCAAGAACACTCAGCCCATCATGGACACAGATGGCTCTT 2279  
Qy 1391 ----- 1390  
Db 2280 ACTTCGTCTACAGCAAGCTCAATGTGCAGAGAGCACTGGGAGCGCAGGAATATCTTCA 2339  
Qy 1391 -----CTCAAGAGAGAAATGTCCCTTCTAAT 1418  
Db 2340 CCTGCTCTGTATCATGAGGCGCTGCACAACCATACTGAGAAGAGCCTCTCCCACT 2399  
Qy 1419 AGTGGGTAGA -----ATTCCCTAGGCGCGCGCAATTA 1453  
Db 2400 CTCCTGGTAAATGATCCCAAGTGTCTTGGAGCCCTCTGTGCTCTCAGCGGCGCGCAATTA 2459  
Qy 1454 ATTCGCTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCT 1513  
Db 2460 ATTCGCTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCT 2519  
Qy 1514 TGTAGATTTAATCAAGAGGATGTGAGATGCCATTTCCCTGAGAGATGAGGCTTCAT 1573  
Db 2520 TGTAGATTTAATCAAGAGGATGTGAGATGCCATTTCCCTGAGAGATGAGGCTTCAT 2579  
Qy 1574 TTTTGTACTTTTATTTGTAACTTATAGTATAGGATTTTGTGCAATTTTGTGTTTC 1633  
Db 2580 TTTTGTACTTTTATTTGTAACTTATAGTATAGGATTTTGTGCAATTTTGTGTTTC 2639  
Qy 1634 TTTCTGTACAGGTTCTCTGTATCAGCCTATCTCCAGCTGATGAATCTTTGTGGTAG 1693

Db 2640 TTCTCGTACGAGCTTGTCTGTATCAGCCTATCTCGCAGCTGATGAATATCTTGTGGTAG 2699  
Qy 1694 GGGTTTGGGAAATCAATTCGAGTTTGTATCTTTTCTTGTATTTCCCACTCTCTTCAGA 1753  
Db 2700 GGGTTTGGGAAATCAATTCGAGTTTGTATCTTTTCTTGTATTTCCCACTCTCTTCAGA 2759  
Qy 1754 GTACAGAAATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATCGGT 1813  
Db 2760 GTACAGAAATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATCGGT 2819  
Qy 1814 AGTTTATCAGAGTAAATTTCTTAACGAGTCAGGACCGGTATGATGAATCTAACATATGG 1873  
Db 2820 AGTTTATCAGAGTAAATTTCTTAACGAGTCAGGACCGGTATGATGAATCTAACATATGG 2879  
Qy 1874 CTGATCGTATCTCCGCAACCGTCACCTGGATGCTGTAGGATAGGCTTGTATGCGG 1933  
Db 2880 CTGATCGTATCTCCGCAACCGTCACCTGGATGCTGTAGGATAGGCTTGTATGCGG 2939  
Qy 1934 GTACTGCGGCGCTCTTTCGGGATATGCTCCATTCGCAAGCATCGCAGTCACTATGGC 1993  
Db 2940 GTACTGCGGCGCTCTTTCGGGATATGCTCCATTCGCAAGCATCGCAGTCACTATGGC 2999  
Qy 1994 GTGCTGTAGCGTATATGCTTGTGCAATTTCTATGCGCACCGTTCCTCGAGCACTG 2053  
Db 3000 GTGCTGTAGCGTATATGCTTGTGCAATTTCTATGCGCACCGTTCCTCGAGCACTG 3059  
Qy 2054 TCCGACCGCTTTCGGCGCGCCAGTCTGCTCGTCTGCTACTTGGAGCACTATCGAC 2113  
Db 3060 TCCGACCGCTTTCGGCGCGCCAGTCTGCTCGTCTGCTACTTGGAGCACTATCGAC 3119  
Qy 2114 TAGCGGATGATGGGACCAACACCGCTCTGCTGATCTATCGAATCTAAATGTAAGTAAA 2173  
Db 3120 TAGCGGATGATGGGACCAACACCGCTCTGCTGATCTATCGAATCTAAATGTAAGTAAA 3179  
Qy 2174 ATCTCTAAATAAATAAGTCCAGTTCCTCATACGAACTTAAACAGCAITTCGGGTG 2233  
Db 3180 ATCTCTAAATAAATAAGTCCAGTTCCTCATACGAACTTAAACAGCAITTCGGGTG 3239  
Qy 2234 AGCATCTAGACCTTCAACAGCAGCAGATCCATCTGCTTGGCCAAATGTTTCAGTCC 2293  
Db 3240 AGCATCTAGACCTTCAACAGCAGCAGATCCATCTGCTTGGCCAAATGTTTCAGTCC 3299  
Qy 2294 CTGAGGAGTTCAGTCTTGTGAAGTGAATCTTCTGGAAGTTCAGTGTAACTCCGCT 2353  
Db 3300 CTGAGGAGTTCAGTCTTGTGAAGTGAATCTTCTGGAAGTTCAGTGTAACTCCGCT 3359  
Qy 2354 GTATTGACGGGCATATCCGTACGTTGGCAAGTGTGGTACCGGAGAGTAATCTCC 2413  
Db 3360 GTATTGACGGGCATATCCGTACGTTGGCAAGTGTGGTACCGGAGAGTAATCTCC 3419  
Qy 2414 ACAACTCTCTGGAGAGTAGGACCAACAAACAGATCCAGCGTGTGTACTTGTATCAAC 2473  
Db 3420 ACAACTCTCTGGAGAGTAGGACCAACAAACAGATCCAGCGTGTGTACTTGTATCAAC 3479  
Qy 2474 ATAGAAGAGCAATTCCTCGATTTTCAGGATCAAGTGTTCAGGAGCGTACTGATGGACAT 2533  
Db 3480 ATAGAAGAGCAATTCCTCGATTTTCAGGATCAAGTGTTCAGGAGCGTACTGATGGACAT 3539  
Qy 2534 TTCAAAAGCCTGCTGTAGTTCGAACCGATAGGTTGTAGAGTGTCAATACACTTGGC 2593  
Db 3540 TTCAAAAGCCTGCTGTAGTTCGAACCGATAGGTTGTAGAGTGTCAATACACTTGGC 3599  
Qy 2594 TACAATTTCAACCTTGGCAACTGCACAGCTTGGTGTGAACAGCACTTCTCAATTCGCG 2653  
Db 3600 TACAATTTCAACCTTGGCAACTGCACAGCTTGGTGTGAACAGCACTTCTCAATTCGCG 3659  
Qy 2654 AAGTCCTTGTCTGTATTCGACAGCAACAGATCACCTGGGAATCAATACCATGTC 2713  
Db 3660 AAGTCCTTGTCTGTATTCGACAGCAACAGATCACCTGGGAATCAATACCATGTC 3719  
Qy 2714 AGCTTGAAGAGAGTCTGAGCAACAGATCTGGATCAGCGTATTTATCAGCAATAAC 2773  
Db 3720 AGCTTGAAGAGAGTCTGAGCAACAGATCTGGATCAGCGTATTTATCAGCAATAAC 3778

2774 TAGAATCTTCAAGAGCCGAGGAGCATGTCAATTAATACTACACAGGCGCTGATGTGTCATTTTG 2833  
3779 TAGAATCTTCAAGAGCCGAGGAGCATGTCAATTAATACTACACAGGCGCTGATGTGTCATTTTG 3838  
2834 AACCATCATCTGGCAGAGTAACGAAGTGTCTTCTGGACCAATATTTTGTACACACTT 2893  
3839 AACCATCATCTGGCAGAGTAACGAAGTGTCTTCTGGACCAATATTTTGTACACACTT 3898  
2894 AGGAACAGTTTCTGTTCCGTAAGCCATAGCACTACTGCTGGGCGCTCTCTGTAGCAC 2953  
3899 AGGAACAGTTTCTGTTCCGTAAGCCATAGCACTACTGCTGGGCGCTCTCTGTAGCAC 3958  
2954 GATACACTTAGCACCAACTTGTGGGCAACGTAGATGACTTCTGGGTAGGGTACCATC 3013  
3959 GATACACTTAGCACCAACTTGTGGGCAACGTAGATGACTTCTGGGTAGGGTACCATC 4018  
3014 CTCTTAGGTGAGATGCAAAACAAATTTCTTTGCAACAGCAACTTTGGCAGGAACACC 3073  
4019 CTCTTAGGTGAGATGCAAAACAAATTTCTTTGCAACAGCAACTTTGGCAGGAACACC 4078  
3074 CAGCATCAGGGAGTGGAGGCAAGTTCGGGTTCCACAGGAATATAGAGGCCAATTT 3133  
4079 CAGCATCAGGGAGTGGAGGCAAGTTCGGGTTCCACAGGAATATAGAGGCCAATTT 4138  
3134 CTCAATAGGTCTTGCAGGAGGAGCAGACTACACAGGCGCAAGTCTCAACTTGCACAGT 3193  
4139 CTCAATAGGTCTTGCAGGAGGAGCAGACTACACAGGCGCAAGTCTCAACTTGCACAGT 4198  
3194 CTCCGTTAGTTAGCTTCATGGAATTCCTGACGTTATCTATAGAGATCAATGGGTCT 3253  
4199 CTCCGTTAGTTAGCTTCATGGAATTCCTGACGTTATCTATAGAGATCAATGGGTCT 4258  
3254 CTCTACGTTATCTGCAATTCGATAAGTTCCTCTGGGAAAGGAGCTCTTAACACAGTGT 3313  
4259 CTCTACGTTATCTGCAATTCGATAAGTTCCTCTGGGAAAGGAGCTCTTAACACAGTGT 4318  
3314 CTCTAACAGGAGCTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTTTGTCAACATTTG 3373  
4319 CTCTAACAGGAGCTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTTTGTCAACATTTG 4378  
3374 AGCAATCTTGCACAACTTGGTTCGACTAATTCANATCTGTTCCGTTTCTGGATAGG 3433  
4379 AGCAATCTTGCACAACTTGGTTCGACTAATTCANATCTGTTCCGTTTCTGGATAGG 4438  
3434 AGCAATCTTGCACAACTTGGTTCGACTAATTCANATCTGTTCCGTTTCTGGATAGG 3493  
4439 AGCAATCTTGCACAACTTGGTTCGACTAATTCANATCTGTTCCGTTTCTGGATAGG 4498  
3494 TTCAATAGACCTTGCAGAGGAGCTTCTTATAGTTAGTTGATCTCTTTAGTGTGTTCTT 3553  
4499 TTCAATAGACCTTGCAGAGGAGCTTCTTATAGTTAGTTGATCTCTTTAGTGTGTTCTT 4558  
3554 GGTGTATCTGCTTGGCATCTCTTCTTCTAGTGACCTTTTAGGACTTCAATCCAG 3613  
4559 GGTGTATCTGCTTGGCATCTCTTCTTCTAGTGACCTTTTAGGACTTCAATCCAG 4618  
3614 GTTCTCTCCACTCTGCTCCAGCTCACACCGTACTTGGCACTCTAATAATGCAAAATA 3673  
4619 GTTCTCTCCACTCTGCTCCAGCTCACACCGTACTTGGCACTCTAATAATGCAAAATA 4678  
3674 AATAAGTCAGACATCTCCAGGCTATCTCTCTGATTTAGTCTTCTGCAAGTTCATC 3733  
4679 AATAAGTCAGACATCTCCAGGCTATCTCTCTGATTTAGTCTTCTGCAAGTTCATC 4738  
3734 AGCTTCTCCCTTAATTTTAGGCTTC - AACAAACTCTGTCGTAATAACCGTTTGGTAT 3792  
4739 AGCTTCTCCCTTAATTTTAGGCTTC - AACAAACTCTGTCGTAATAACCGTTTGGTAT 4798  
3793 AAGAACCTTCTGGAGATGCTTACGATCCACAGGTCGCTCCATGCTCTAAGAC 3852  
4799 AAGAACCTTCTGGAGATGCTTACGATCCACAGGTCGCTCCATGCTCTAAGAC 4857

3853 CTTTTGATTTGGCCAAAACAGGAGTGGTTCCAGTGTACAGAAACCAACACCTGTTTCTT 3912  
4858 CTTTTGATTTGGCCAAAACAGGAGTGGTTCCAGTGTACAGAAACCAACACCTGTTTCTT 4917  
3913 CAACACAAAATTTCAAGCAGTCTCCATCACAAATCCAAATTCGATACCCAGCAACTTTTCAAG 3972  
4918 CAACACAAAATTTCAAGCAGTCTCCATCACAAATCCAAATTCGATACCCAGCAACTTTTCAAG 4977  
3973 TTGCTCCAGATGTAGACCTTTATACCAAAACCTGTGACGAGGATTTGTAGACTCCAG 4032  
4978 TTGCTCCAGATGTAGACCTTTATACCAAAACCTGTGACGAGATTTGTAGACTCCAG 5037  
4033 TTGCTGCTCTTATAGCTCCGGAATAGACTTTTTCGAGGAGTACACAGGCGCCCAACGAGT 4092  
5038 TTGCTGCTCTTATAGCTCCGGAATAGACTTTTTCGAGGAGTACACAGGCGCCCAACGAGT 5097  
4093 AATTAGAGAGTCAGCCACCAAAAGTAGTGAATAGACCATTCGGGGCGGTCACTAGTCAAG 4152  
5098 AATTAGAGAGTCAGCCACCAAAAGTAGTGAATAGACCATTCGGGGCGGTCACTAGTCAAG 5157  
4153 AGCCAAACAAAATTTCACTGACAGGAACTTTTTCGACATCTTCAGAAAAGTTCGTATTTCAG 4212  
5158 AGCCAAACAAAATTTCACTGACAGGAACTTTTTCGACATCTTCAGAAAAGTTCGTATTTCAG 5217  
4213 TAGTCAATTTGCCAGCATCAATTAATGGGATTTATACAGAGCAACAGTGGAGTCAAT 4272  
5218 TAGTCAATTTGCCAGCATCAATTAATGGGATTTATACAGAGCAACAGTGGAGTCAAT 5277  
4273 CTACCAACTTTGGGCTCTCAGAAAGCATAAACAGTTCCTACTCCGCCATTAGTGAAC 4332  
5278 CTACCAACTTTGGGCTCTCAGAAAGCATAAACAGTTCCTACTCCGCCATTAGTGAAC 5337  
4333 TTTTCAAAATCGCCAGTCGAGTGGAGAAAGGCAACAGTACTAGCATTTAGCGGCAAGG 4392  
5338 TTTTCAAAATCGCCAGTCGAGTGGAGAAAGGCAACAGTACTAGCATTTAGCGGCAAGG 5397  
4393 ATGCAACTTTTATCAACAGGTCCTATAGATACCTAGCGCTGGGATCTCTTTGGA 4452  
5398 ATGCAACTTTTATCAACAGGTCCTATAGATACCTAGCGCTGGGATCTCTTTGGA 5457  
4453 CAATCTTTCTGCCAAATCTAGGTCGAAATCACTTTCATTTGATACCAATTTAGTCAACT 4512  
5458 CAATCTTTCTGCCAAATCTAGGTCGAAATCACTTTCATTTGATACCAATTTAGTCAACT 5517  
4513 TGAGCAAGTTGTGATCTAGTCTCTCAAAATTTGGTCTCTGTAAACGATGACTCAACTTGA 4572  
5518 TGAGCAAGTTGTGATCTAGTCTCTCAAAATTTGGTCTCTGTAAACGATGACTCAACTTGA 5577  
4573 CATTAACCTTGAAGCTCAGTCGATTCAGTGAACCTTATCAGTGAACCTTATCAAACTCTGCAACACTTG 4632  
5578 CATTAACCTTGAAGCTCAGTCGATTCAGTGAACCTTATCAGTGAACCTTATCAAACTCTGCAACACTTG 5637  
4633 CATAGGAAACACGCTTTCTCTACCAAACTCAAGAAATTTCAAACTCTGCAACACTTG 4692  
5638 CATAGGAAACACGCTTTCTCTACCAAACTCAAGAAATTTCAAACTCTGCAACACTTG 5697  
4693 CGTATCAGGTAGCAAGGAAATGTCATCTTGAAGTGGAGTGGAGTGTCTTTCAG 4752  
5698 CGTATCAGGTAGCAAGGAAATGTCATCTTGAAGTGGAGTGGAGTGTCTTTCAG 5757  
4753 AAATCTGAAAGCGTATTTTATATCAGTGAAGTCACTCAGGAGATCTCTACGCG 4812  
5758 AAATCTGAAAGCGTATTTTATATCAGTGAAGTCACTCAGGAGATCTCTACGCG 5817  
4813 GACGATCTGTGCGGACCTGAGTTCGATCAACCGGCGCACAGTGGTGTGCGG 4872  
5818 GACGATCTGTGCGGACCTGAGTTCGATCAACCGGCGCACAGTGGTGTGCGG 5877  
4873 CCTATATCCCGACATCAACCGATGGGAGATCGGGCTCGCCACTTCGGGCTCATGAGCG 4932  
5878 CCTATATCCCGACATCAACCGATGGGAGATCGGGCTCGCCACTTCGGGCTCATGAGCG 5937  
4933 CTTGTTTCGGCTGGGTATGTTGGCAGGCGCGCTGGGAGGACTGTTGGCGGCACTCT 4992

Db 5938 CTTGTTTCGGCTGGGTATGTTGGCAGGCCCGTGGCGGGGACTGTTGGCGCATCT 5997  
QY CTTTGCATGCAC----- 5004  
Db 5998 CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 6057  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6058 CATTGCACAAGATAAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACATAAACAG 6117  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6118 TAATACAAGGGTGTATGAGCCATATTCAACGGGAAGCGTCTTGCTCAAGSCCGGATT 6177  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6178 AAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGGATATGTCGGCA 6237  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6238 ATCAGTGGCACAATCTATTCGATTGTATGGGAAGCCGATGCGCAGAGTTGTTTCTGAA 6297  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6298 ACATGCAAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAACTGGCT 6357  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6358 GACGGAATTTATGCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGTGATGATG 6417  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6418 GTTACTCACCACTGGATCCCGGGGAAACAGCATTCAGGTATTAGAGAATATCCTGA 6477  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6478 TTCAAGTGAATAATTGTTGATGCGCTGGCAGTGTTCCTGCGCGTGTGCAATTCGATCC 6537  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6538 TGTTTGAATTGCTCTTTAAACAGCGATCGCGTATTTGCTGCTCAGGCGCAATCAG 6597  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6598 AATGAATAACGGTTTGGTTGATGCGAGTGATTTGATGACGAGCGTAATGGCTGGCTGT 6657  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6658 TGAACAAGTCTGGAAGAAATGCAATAGCTTTTGGCAATTCACCGGATTCAGTCGTAC 6717  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6718 TCATGGTGATTTCTCACTTGTATAACCTTATTTTTCAGGAGGGAATTAATAGGTTGTAT 6777  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6778 TGATGTTGACGAGTCGGAATCGCAGACGATACCGAGATCTTGCCATCCTATGGAACGT 6837  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6838 CCTCGTGAGTTTCTCCTTCATTACAGAAACGGCTTTTCAAATAATGTTATGTATAA 6897  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6898 TCCTGATATGAATAAATGCGATTTTCATTGTGATGCTCGATGAGTTTCTTAATCAGAAAT 6957  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 6958 GGTTAATTGGTTGTAACACTGGCAGAGCATTAAGTGAATTGACGGGACGGCGCTTGT 7017  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004

Db 7018 TGAATAAATCGAACTTTTGTGCTGAGTTGAAGGATCAGATCAGCATCTTCCCGACACGCA 7077  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7078 GACGTTCCGTGGCAAGCAAAAGTTCAAATCACCACCTGCTCCACCTACACAAAGCT 7137  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7138 CTATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGGATTCAGGCTGGTAT 7197  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7198 GAGTCAGCAACACCTTTTTCACGAGGCGAGCTCAGCGCCCCCCCCCTCGAGGTCCC 7257  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7258 ACAGCGCGGTGCTCAACGGCCTCAACCTACTACTGCGCTGCTTCCCTAAATGCAAGGATCG 7317  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7318 CATGAAGGAGAGCGTCAAGTATCTATGATTGGAAGTATGGAATGGTATGATACCGGATTC 5132  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7319 CATGAAGGAGAGCGTCAAGTATCTATGATTGGAAGTATGGAATGGTATGATACCGGATTC 7377  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7378 TTCAAGTGTCTGAGTCTCTATCAGATTATGCCCACTAAAGCAACCGGAGGAGAT 7437  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7438 TTCAAGTGTCTGAGTCTCTATGATTGGAAGTATGGAATGGTATGATACCGGATTC 7497  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7498 GTTATGACAGCAGAAATGTCTCTTGGAGACAGTAAATGAAGTCCCAACCAATAAGAAA 7557  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7558 TCCTTGTATCAGAAACAACTTCTTGTGTTTGGAACTTTTTCGGTGCTTGAATATAAA 7617  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7618 TGTAAGTGGATATGTCGGGTAGGAATGGAGCGGCAAAATGCTTACCTTTCGACCTTCA 5432  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7619 TGTAAGTGGATATGTCGGGTAGGAATGGAGCGGCAAAATGCTTACCTTTCGACCTTCA 7677  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7678 AGAGGTATGAGGTTTGTAGATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 5492  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7679 AGAGGTATGAGGTTTGTAGATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 7737  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7738 TCAAAACCATTCGGAATCCAGAGAAATCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5552  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7739 TCAAAACCATTCGGAATCCAGAGAAATCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7797  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7798 GCGGTCTTGAACCTGACAAATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7857  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7859 GCGGTCTTGAACCTGACAAATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5612  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7860 GCGGTCTTGAACCTGACAAATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7857  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7861 CTAGTCTTGTATCTAAATATCTTGACGAGCGGAGGATAATACCAATCTTAAACT 5672  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7862 CTAGTCTTGTATCTAAATATCTTGACGAGCGGAGGATAATACCAATCTTAAACT 7917  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7863 CTAGTCTTGTATCTAAATATCTTGACGAGCGGAGGATAATACCAATCTTAAACT 5732  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7864 CTAGTCTTGTATCTAAATATCTTGACGAGCGGAGGATAATACCAATCTTAAACT 7977  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7865 GTCTGATCTCATCAACTGAGGGGACTATCTTGTGTTTGTAGAGAAATTCGGGAGATGCG 5792  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7866 GTCTGATCTCATCAACTGAGGGGACTATCTTGTGTTTGTAGAGAAATTCGGGAGATGCG 8037  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7867 ATATCGAGAAAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAAGATCTCTGCTCG 5852  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7868 ATATCGAGAAAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAAGATCTCTGCTCG 8097  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7869 CGCGTTTCGGTGTAGCGGTGAAAACTCTTGACACATGACAGTCCCGGAGACGCTCACAG 5912  
QY CTTTGGACCTGCAGGGGGGGGGGAAAGCACCAGCTTGTTGCTCAAATCTCTGATGTTA 5004  
Db 7870 CGCGTTTCGGTGTAGCGGTGAAAACTCTTGACACATGACAGTCCCGGAGACGCTCACAG 8157

QY 5913 CTTGTCCTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGCTTG 5972  
 Db 8158 CTTGTCCTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGCTTG 8217  
 QY 5973 GCGGCTGTCGGGCGCAGCCATGACCCAGTCACCTAGCGATAGCGGAGTGTATCTGGCT 6032  
 Db 8218 GCGGCTGTCGGGCGCAGCCATGACCCAGTCACCTAGCGATAGCGGAGTGTATCTGGCT 8277  
 QY 6033 TAACTATCGGCGATCAGAGCGAGATTGTACTGAGAGTGACCATATGCGGTGTGAATACC 6092  
 Db 8278 TAACTATCGGCGATCAGAGCGAGATTGTACTGAGAGTGACCATATGCGGTGTGAATACC 8337  
 QY 6093 GCACAGATCGGTAAAGGAGAAATACCGCATCAGGCGCTCTCCGCTTCTCGCTCACTGA 6152  
 Db 8338 GCACAGATCGGTAAAGGAGAAATACCGCATCAGGCGCTCTCCGCTTCTCGCTCACTGA 8397  
 QY 6153 CTCGCTGCGCTCGGTCTGCGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAA 6212  
 Db 8398 CTCGCTGCGCTCGGTCTGCGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAA 8457  
 QY 6213 ACGGTTATCCAGAAATCAGGGGTAACGCGGAGAAACATGTGACGAAAGGCCAGCA 6272  
 Db 8458 ACGGTTATCCAGAAATCAGGGGTAACGCGGAGAAACATGTGACGAAAGGCCAGCA 8517  
 QY 6273 AAAGGCCAGGAACCGTAAAGGCGCGCTTGCTGCGCTTTTCCATAGGCTCCGCCCCC 6332  
 Db 8518 AAAGGCCAGGAACCGTAAAGGCGCGCTTGCTGCGCTTTTCCATAGGCTCCGCCCCC 8577  
 QY 6333 TGACGAGATCAAAAATCAGAGCTCAAGTCAGAGGTGGGGAACCGCAGAGACTATA 6392  
 Db 8578 TGACGAGATCAAAAATCAGAGCTCAAGTCAGAGGTGGGGAACCGCAGAGACTATA 8637  
 QY 6393 AAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTGTTCGAGACCTGCC 6452  
 Db 8638 AAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTGTTCGAGACCTGCC 8697  
 QY 6453 GCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCTC 6512  
 Db 8698 GCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCTC 8757  
 QY 6513 ACGCTGAGGATCTCAGTTCGGTGTAGTGTCTCGCTCCAGCTGGGCTGTGTGACGA 6572  
 Db 8758 ACGCTGAGGATCTCAGTTCGGTGTAGTGTCTCGCTCCAGCTGGGCTGTGTGACGA 8817  
 QY 6573 ACCCCCGCTTACGCGCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCAACCC 6632  
 Db 8818 ACCCCCGCTTACGCGCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCAACCC 8877  
 QY 6633 GGTAAACAGACTATTCGCTACCGCTGGCAGCAGCCTGTTACAGGATAGCAGAGCGAG 6692  
 Db 8878 GGTAAACAGACTATTCGCTACCGCTGGCAGCAGCCTGTTACAGGATAGCAGAGCGAG 8937  
 QY 6693 GTATGAGCGGTGTACAGAGTTCCTGAAGTGGTGGCTTAACCTACGCGCTACCTAGAAG 6752  
 Db 8938 GTATGAGCGGTGTACAGAGTTCCTGAAGTGGTGGCTTAACCTACGCGCTACCTAGAAG 8997  
 QY 6753 GACAGTATTTGGTATCTCGGCTCTGCTGAAGCGAGTTACCTTCGGAAGAGAGTTGGTAG 6812  
 Db 8998 GACAGTATTTGGTATCTCGGCTCTGCTGAAGCGAGTTACCTTCGGAAGAGAGTTGGTAG 9057  
 QY 6813 CTCCTGATCCGCAAAACCAACCGCTGCTAGCGGTGGTGTGTTTGTGTTGCAAGCAGCA 6872  
 Db 9058 CTCCTGATCCGCAAAACCAACCGCTGCTAGCGGTGGTGTGTTTGTGTTGCAAGCAGCA 9117  
 QY 6873 GATTACCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATCTCTTCTAGCGGTCTGA 6932  
 Db 9118 GATTACCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATCTCTTCTAGCGGTCTGA 9177  
 QY 6933 CGCTCAGTGAAGCAAACTCAGTTTAAAGGATTTTGGTCATGAGATTATCAAAAGGAT 6992  
 Db 9178 CGCTCAGTGAAGCAAACTCAGTTTAAAGGATTTTGGTCATGAGATTATCAAAAGGAT 9237

QY 6993 CTTACCTAGATCCCTTTTAAATTTAAATGAAGTTTAAATCAATCTAAAGTATATATGA 7052  
 Db 9238 CTTACCTAGATCCCTTTTAAATTTAAATGAAGTTTAAATCAATCTAAAGTATATATGA 9297  
 QY 7053 GTAAATCTGGTGTGACAGTACCAATGCTTAATCAGTGAGGCACTATCTCAGGATCTG 7112  
 Db 9298 GTAAATCTGGTGTGACAGTACCAATGCTTAATCAGTGAGGCACTATCTCAGGATCTG 9357  
 QY 7113 TCTATTCTGCTTCATCCATAGTTCCTGACTCCCGCTGCTGTAGATAACTACGATACGGA 7172  
 Db 9358 TCTATTCTGCTTCATCCATAGTTCCTGACTCCCGCTGCTGTAGATAACTACGATACGGA 9417  
 QY 7173 GGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCC 7232  
 Db 9418 GGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCC 9477  
 QY 7233 AGATTTATCAGCAATAAACCAGCCAGCGGAGGCGGAGCGAGAGTGTGCTGCAAC 7292  
 Db 9478 AGATTTATCAGCAATAAACCAGCCAGCGGAGGCGGAGCGAGAGTGTGCTGCAAC 9537  
 QY 7293 TTTATCCGCTCCATCCAGTCTTAAATGTTTTCGGGGAAGCTAGAGTAAGTTCGCC 7352  
 Db 9538 TTTATCCGCTCCATCCAGTCTTAAATGTTTTCGGGGAAGCTAGAGTAAGTTCGCC 9597  
 QY 7353 AGTTAATAGTTTGGCAACGTTTGGCCATTCCTGAGGATCGTGGTGTACGCTCGTC 7412  
 Db 9598 AGTTAATAGTTTGGCAACGTTTGGCCATTCCTGAGGATCGTGGTGTACGCTCGTC 9657  
 QY 7413 GTTTGATAGTGTTCATTCAGCTCCGCTTCCCAACGATCAAGCGGAGTTACATGATCCCC 7472  
 Db 9658 GTTTGATAGTGTTCATTCAGCTCCGCTTCCCAACGATCAAGCGGAGTTACATGATCCCC 9717  
 QY 7473 CATGTTGTCAAAAAGCGGTAGCTCTTCGGTCTCCGATCGTGTGTCAGAGTAAGTT 7532  
 Db 9718 CATGTTGTCAAAAAGCGGTAGCTCTTCGGTCTCCGATCGTGTGTCAGAGTAAGTT 9777  
 QY 7533 GGGCGAGTGTATCACTCATGTTATGCGAGCACTGCATATTTCTTACTGTCTATGCC 7592  
 Db 9778 GGGCGAGTGTATCACTCATGTTATGCGAGCACTGCATATTTCTTACTGTCTATGCC 9837  
 QY 7593 ATCCGTAAGATGCTTTCTGAGCTGTGAGTATCAACCAAGTCACTCTGAGATAGTG 7652  
 Db 9838 ATCCGTAAGATGCTTTCTGAGCTGTGAGTATCAACCAAGTCACTCTGAGATAGTG 9897  
 QY 7653 TATGCGCGCAGCGAGTGTCTTTCGGCGCGTCAACACGGGATAATACCGCGCCACATAG 7712  
 Db 9898 TATGCGCGCAGCGAGTGTCTTTCGGCGCGTCAACACGGGATAATACCGCGCCACATAG 9957  
 QY 7713 CAGAACTTTAAAGTCTCATCATTTGGAAGAAAGTTCCTCGGGCGGAAACTCTCAAGAT 7772  
 Db 9958 CAGAACTTTAAAGTCTCATCATTTGGAAGAAAGTTCCTCGGGCGGAAACTCTCAAGAT 10017  
 QY 7773 CTTACCGCTGTTCAGATCCAGTTCGATTAACCACTCGTGACCCCACTGATCTTCAGC 7832  
 Db 10018 CTTACCGCTGTTCAGATCCAGTTCGATTAACCACTCGTGACCCCACTGATCTTCAGC 10077  
 QY 7833 ATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA 7892  
 Db 10078 ATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA 10137  
 QY 7893 AAGGGAATAGGGCGCAGCGGAAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCTCT 7952  
 Db 10138 AAGGGAATAGGGCGCAGCGGAAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCTCT 10197  
 QY 7953 TTGAAGCATTTTACAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGATTTAGAA 8012  
 Db 10198 TTGAAGCATTTTACAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGATTTAGAA 10257  
 QY 8013 AATTAACAAATAGGGGTTCCGGCGACATTTCCCGGAAAGTCCACCTGAGCTCTAAGA 8072  
 Db 10258 AATTAACAAATAGGGGTTCCGGCGACATTTCCCGGAAAGTCCACCTGAGCTCTAAGA 10317  
 QY 8073 AACCATTTTATCATGACATTAACCTATAAAATAGGCGGTATCAGGAGGCCCTTTCTGCT 8132

Db	10318	AACCAATTATTTCATGACATTAACTATAAAATAGCGTTATCAGAGGCCCTTTCGTCT	10377
Qy	8133	TCAGAATAATTCTCATGTTTGACAGCTTATCATCGATAAGCTGACTCATGTTGGTATT	8192
Db	10378	TCAGAATAATTCTCATGTTTGACAGCTTATCATCGATAAGCTGACTCATGTTGGTATT	10437
Qy	8193	GTGAAATAGACGACATCGGAACACTGAAAAATAACAGTTATTATTTCG	8241
Db	10438	GTGAAATAGACGACATCGGAACACTGAAAAATAACAGTTATTATTTCG	10486

Search completed: September 24, 2004, 04:29:51  
Job time : 21002.8 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 04:29:55 ; Search time 1464.26 Seconds  
(without alignments)  
17786.465 Million cell updates/sec

Title: US-10-030-390-1

Perfect score: 5142

Sequence: 1 gaattcagtcgtaccg.....ccgtctgtaagtccttaag 5142

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4232.8	82.3	10929	15	US-10-032-393-7
2	1493	29.0	5865	13	US-10-402-466A-2
3	1493	29.0	6694	13	US-10-402-466A-1
4	1493	29.0	7336	15	US-10-462-128-3
5	1493	29.0	7336	15	US-10-084-814-3
6	1493	29.0	7336	16	US-10-442-013-4
7	1493	29.0	8198	13	US-10-402-466A-7
8	1493	29.0	9286	13	US-10-402-466A-10
9	1487.6	28.9	5302	16	US-10-385-415-2
10	1487.6	28.9	5767	16	US-10-385-415-4
11	1207.6	23.5	5793	10	US-09-869-855A-3
12	1207.6	23.5	5943	10	US-09-869-855A-2
13	981.2	19.1	6852	15	US-10-032-393-16
14	981.2	19.1	6852	15	US-10-032-393-53

C 15	978.4	19.0	3867	13	US-10-466-642-24	Sequence 24, Appl
C 16	978.4	19.0	3867	17	US-10-466-717-24	Sequence 24, Appl
C 17	978	19.0	6725	13	US-10-466-642-22	Sequence 22, Appl
C 18	978	19.0	6725	17	US-10-466-717-22	Sequence 22, Appl
C 19	978	19.0	6806	13	US-10-466-642-23	Sequence 23, Appl
C 20	978	19.0	6806	17	US-10-466-717-23	Sequence 23, Appl
C 21	978	19.0	6886	13	US-10-466-642-20	Sequence 20, Appl
C 22	978	19.0	6886	17	US-10-466-717-20	Sequence 20, Appl
C 23	978	19.0	7140	13	US-10-466-642-21	Sequence 21, Appl
C 24	978	19.0	7140	17	US-10-466-717-21	Sequence 21, Appl
C 25	978	19.0	8320	9	US-09-813-453A-71	Sequence 71, Appl
C 26	978	19.0	8803	17	US-10-466-717-27	Sequence 27, Appl
C 27	977	19.0	7701	17	US-10-466-641-25	Sequence 25, Appl
C 28	977	19.0	7926	17	US-10-466-641-24	Sequence 24, Appl
C 29	943	18.3	10186	15	US-10-234-329-3	Sequence 3, Appl
C 30	867	16.9	3156	13	US-10-681-086-1	Sequence 1, Appl
C 31	866.8	16.9	2517	9	US-09-893-600-1	Sequence 1, Appl
C 32	866.8	16.9	2517	9	US-09-893-499-1	Sequence 1, Appl
C 33	866.8	16.9	2517	9	US-09-894-030-1	Sequence 1, Appl
C 34	771	15.0	6837	10	US-09-928-847B-49	Sequence 49, Appl
C 35	759	14.8	759	17	US-10-718-628-10	Sequence 10, Appl
C 36	692	13.5	692	9	US-09-974-300-6910	Sequence 6910, Ap
C 37	541	10.5	6405	10	US-09-869-855A-1	Sequence 1, Appl
C 38	347.2	6.8	8704	13	US-10-466-642-25	Sequence 25, Appl
C 39	347.2	6.8	8704	17	US-10-466-717-25	Sequence 25, Appl
C 40	347.2	6.8	9292	9	US-09-813-453A-74	Sequence 74, Appl
C 41	347.2	6.8	9396	9	US-09-813-453A-73	Sequence 73, Appl
C 42	198.6	3.9	447	16	US-10-166-349-11	Sequence 11, Appl
C 43	198.6	3.9	447	16	US-10-166-349-19	Sequence 19, Appl
C 44	170.2	3.3	4914	16	US-10-310-740-11	Sequence 11, Appl
C 45	162	3.2	5502	9	US-09-736-457-785	Sequence 785, App

ALIGNMENTS

RESULT 1

US-10-032-393-7  
; Sequence 7, Application US/10032393  
; Publication No. US20030027286A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Wall, Daniel  
; APPLICANT: Gross, Molly  
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
; FILE REFERENCE: ELITRA.010A  
; CURRENT APPLICATION NUMBER: US/10/032,393  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR FILING DATE: 2001-09-06  
; PRIOR FILING DATE: 2001-09-06  
; PRIOR FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 10929  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Vector pEP25  
US-10-032-393-7

Query Match 82.3%; Score 4232.8; DB 15; Length 10929;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 4323; Conservative 0; Mismatches 17; Indels 10; Gaps 7;  
Qy 651 CTGATTTTCATTTTTCATTTCTCAAACTGCATATCATATGTAATCGCTCTCTTTT 710  
Db 6491 CTGATTTTCATTTTTCATTTCTCAAACTGCATATCATATGTAATCGCTCTCTTTT 6550  
Qy 711 GGTGGCAAAATGTGAGGCATTTTCGCTCTTTTCGGCGAGGCTAGTTACCCCTTAAGTTAT 770



|||||  
6551 GGTGGCAAAATGAGAGCAATTCGCTCTTCGCGAGGCTAGTTACCCCTTAAGTTAT 6610  
|||||  
771 TGGTATGACTGGTTTAAAGCCAAAAAAGTTGGCTTTTGGTACCTATTAATGATCGTT 830  
|||||  
6611 TGGTATGACTGGTTTAAAGCCAAAAAAGTTGGCTTTTGGTACCTATTAATGATCG-- 6668  
|||||  
831 TTGAGAAACCGACTGTAARAAAGTACAGTCGGCATTTCTCATATTAATAAGCCAGTCAT 890  
|||||  
6669 TTGAGAAACCGACTGTAARAAAGTACAGTCGGCATTTCTCATATTAATAAGCCAGTCAT 6728  
|||||  
891 TAGGCTTATCTGACAAATTCCTGAAATAGAGTTTCAATAAACAATTCCTGCAATAC 950  
6729 TAGGCTTATCTGACAAATTCCTGAAATAGAGTTTCAATAAACAATTCCTGCAATAC 6788  
|||||  
951 AAAACGAATGATGTACCTGTAAGATAGCGGTAAATATTTGAATTAACCTTTTAAATGA 1010  
6789 AAAACGAATGATGTACCTGTAAGATAGCGGTAAATATTTGAATTAACCTTTTAAATGA 6848  
|||||  
1011 ATTTTCCTGCTGTAATAATGGTGAAGTAAATTAATTAATTAATTAATTAATTAATTA 1070  
6849 ATTTTCCTGCTGTAATAATGGTGAAGTAAATTAATTAATTAATTAATTAATTAATTA 6908  
|||||  
1071 ACCCAGTAAATGAAGTCATGGAATATAAGAAAGAAAGAAAGCAATTTTCAGGTATAGGTG 1130  
6909 ACCCAGTAAATGAAGTCATGGAATATAAGAAAGAAAGAAAGCAATTTTCAGGTATAGGTG 6968  
|||||  
1131 TTTTGGGAAACAATTTCCCGAACCATTTATTTCTCTACATCAGAAGGTATATAATCAT 1190  
6969 TTTTGGGAAACAATTTCCCGAACCATTTATTTCTCTACATCAGAAGGTATATAATCAT 7028  
|||||  
1191 AAACTCTTTGAAGTCATTTTACAGGAGTCGCAAAATACAGAGAAATGTTTTAGATACAC 1250  
7029 AAACTCTTTGAAGTCATTTTACAGGAGTCGCAAAATACAGAGAAATGTTTTAGATACAC 7088  
|||||  
1251 CATCAAAAATTTGATAAAGTGGCTTAACTTTTCCCAATACCTTAACTCCGTCGGTAT 1310  
7089 CATCAAAAATTTGATAAAGTGGCTTAACTTTTCCCAATACCTTAACTCCGTCGGTAT 7148  
|||||  
1311 TGTAACCAAGTCTTAAAGCTGATTTGAGTTTATCACCCCTTGTCACCTAAAGAAATAAATG 1370  
7149 TGTAACCAAGTCTTAAAGCTGATTTGAGTTTATCACCCCTTGTCACCTAAAGAAATAAATG 7208  
|||||  
1371 CAGGCTAAATTTATATCCTTCTGTTTATGTTTCGTTATATAAACAATAATCAATTT 1430  
7209 CAGGCTAAATTTATATCCTTCTGTTTATGTTTCGTTATATAAACAATAATCAATTT 7268  
|||||  
1431 CTGTGGTTTACTAAAGTCGTTTGGTTCAAAATATGATTAATATCTCTTTTCTCT 1490  
7269 CTGTGGTTTACTAAAGTCGTTTGGTTCAAAATATGATTAATATCTCTTTTCTCT 7328  
|||||  
1491 TCCAAATGCTAAATCAATTTTAAAGTTCAATTTGATATGCTTAAATTTTATCT 1550  
7329 TCCAAATGCTAAATCAATTTTAAAGTTCAATTTGATATGCTTAAATTTTATCT 7388  
|||||  
1551 AAAGTGAATTTAGAGGCTTACTTGTCTTCTTCAATTAAGATCAATCTCTTTTAA 1610  
7389 AAAGTGAATTTAGAGGCTTACTTGTCTTCTTCAATTAAGATCAATCTCTTTTAA 7448  
|||||  
1611 AGTCAATATTAAGTGAATTAATATAATTTTAAATAATFCCCACTTTTCCAAATTTTC 1670  
7449 AGTCAATATTAAGTGAATTAATATAATTTTAAATAATFCCCACTTTTCCAAATTTTC 7508  
|||||  
1671 GTTTGTTGAATTAAGGCTTCTAGTTGAAGAT-AAAAGACCAATTAATAAATGTTG 1729  
7509 GTTTGTTGAATTAAGGCTTCTAGTTGAAGATTAATAAAGACCAATTAATAAATGTTG 7568  
|||||  
1730 CTTTGTGCTTTTAAAGGATTTGAGCGTAGCGGAAATTCCTTTTCTTTCTTATCTTGA 1789  
7569 CTTTGTGCTTTTAAAGGATTTGAGCGTAGCGGAAATTCCTTTTCTTTCTTATCTTGA 7628  
|||||  
1790 TAAATAGGATTAATTCGGGATAGACTGTAACATCTTCACGCATAAATCCCTTTTC 1849

7629 TAAATAGGATTAATTCGGGATAGACTGTAACATTTCTCACGCATAAATCCCTTTTC 7688  
QY  
1850 ATTTTCTAATGTAATCTATTACCTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1909  
Db  
7689 ATTTTCTAATGTAATCTATTACCTTATTAATTAATTAATTAATTAATTAATTAATTAAT 7748  
QY  
1910 TTCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTTTCGGTTAATTCGCCCATGA 1969  
7749 TTCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTTTCGGTTAATTCGCCCATGA 7808  
Db  
1970 CAGCCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2029  
7809 CAGCCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7868  
QY  
2030 AATTTATTAGAGTCAATCGTTCAAAATGGTATGCGTTTTGACACATCCACTATATATCCGT 2089  
7869 AATTTATTAGAGTCAATCGTTCAAAATGGTATGCGTTTTGACACATCCACTATATATCCGT 7928  
Db  
2090 GTCGTTCTGTCACCTCCGAAATCCCAATTCAGAAATTTCTTAGCGATTTCCAGAAAGTTTCT 2149  
7929 GTCGTTCTGTCACCTCCGAAATCCCAATTCAGAAATTTCTTAGCGATTTCCAGAAAGTTTCT 7988  
QY  
2150 CAGAGTCGGAAGTTGACCGACATTTACGAACTGSCACAGATGGTCATAAACCCTGAAGAA 2209  
7989 CAGAGTCGGAAGTTGACCGACATTTACGAACTGSCACAGATGGTCATAAACCCTGAAGAA 8048  
Db  
2210 GATCTGATTTGCTTAACTGCTTCAAGTAAAGCCGAGCGCTCTGCTATATAACAGATGGAT 2269  
8049 GATCTGATTTGCTTAACTGCTTCAAGTAAAGCCGAGCGCTCTGCTATATAACAGATGGAT 8108  
QY  
2270 GATCGAGACCAATCAACATGCGACCTGCGATTTGCTTACCTGTACAGTCAAGATGGTAGAA 2329  
8109 GATCGAGACCAATCAACATGCGACCTGCGATTTGCTTACCTGTACAGTCAAGATGGTAGAA 8168  
Db  
2330 ATGTTGTGCGTCTTGCCACACGAATATTACGCCATTTGCTTGCATATTTCAAAACAGCTCTT 2389  
8169 ATGTTGTGCGTCTTGCCACACGAATATTACGCCATTTGCTTGCATATTTCAAAACAGCTCTT 8228  
QY  
2390 CTACGATTAAGGCGCAAAATCGCATCGTGAGACGTTTGGGCTTCTACCGATTTAGCAGTTT 2449  
8229 CTACGATTAAGGCGCAAAATCGCATCGTGAGACGTTTGGGCTTCTACCGATTTAGCAGTTT 8288  
Db  
2450 GATACACTTTCTTAAGTATCCACCTGAATCATTAATCGGCAAAATAGAGAAATTTGAC 2509  
8289 GATACACTTTCTTAAGTATCCACCTGAATCATTAATCGGCAAAATAGAGAAATTTGAC 8348  
QY  
2510 CATGTTAAGCGGCGCAATCTGATTTCCACTGAGATGCGATAATCTAGTAGAATCTCTTCCG 2569  
8349 CATGTTAAGCGGCGCAATCTGATTTCCACTGAGATGCGATAATCTAGTAGAATCTCTTCCG 8408  
Db  
2570 TATCAAAATTCATTTCCACCTTCCACTCACCGTTGCTCCATTTCAATGCTGCTGCTTCTGCTT 2629  
8409 TATCAAAATTCATTTCCACCTTCCACTCACCGTTGCTCCATTTCAATGCTGCTGCTTCTGCTT 8468  
QY  
2630 CCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGCGGCCCATCAGTCTGACGAC 2689  
8469 CCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGCGGCCCATCAGTCTGACGAC 8528  
Db  
2690 CAGAGAGCCATAAACACCAATAGCCTTAAACATCATCTCCCATATTTATCCCAATATTCGTT 2749  
8529 CAGAGAGCCATAAACACCAATAGCCTTAAACATCATCTCCCATATTTATCCCAATATTCGTT 8588  
QY  
2750 CTTTAAATTCATGAACAATCTTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2809  
8589 CTTTAAATTCATGAACAATCTTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 8648  
Db  
2810 CTATTTCTCATTTCCCTTTTTCAGATAATTTTTCAGTTTCTTCTTCTTCTTCTTCTTCTTCTT 2869  
8649 CTATTTCTCATTTCCCTTTTTCAGATAATTTTTCAGTTTCTTCTTCTTCTTCTTCTTCTTCTT 8708  
QY  
2870 GAGCACCGTTCTTATTTAGCTATTAATAACTCGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2929  
8709 GAGCACCGTTCTTATTTAGCTATTAATAACTCGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 8768



GENERAL INFORMATION:  
APPLICANT: Park, Sukjoon  
APPLICANT: Giri, Lallan  
TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING  
AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS  
FILE REFERENCE: 18933-00005  
CURRENT APPLICATION NUMBER: US/10/402,466A  
CURRENT FILING DATE: 2003-03-20  
PRIOR APPLICATION NUMBER: 60/372,152  
PRIOR FILING DATE: 2002-04-12  
NUMBER OF SEQ ID NOS: 34  
SEQ ID NO 2  
LENGTH: 5865  
TYPE: DNA  
ORGANISM: Artificial  
FEATURE:  
NAME/KEY:  
LOCATION:  
OTHER INFORMATION: Artificial DNA sequence to be used as one of the two backbones of  
E. coli-Bacillus anthracis shuttle vectors. Designated as pbp II  
US-10-402-466A-2

Query Match 29.0%; Score 1493; DB 13; Length 5865;  
Best Local Similarity 98.1%; Pred. No. 2.7e-282;  
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY	1778	TTCTTATCTTGATAATAAGGTTAACTATTGCGGGATAGACTGTAAACATTTCTCAGCGATA	1837
DB	5526	TACTCTTTTAAATCCCGACTGCGAATGCGGGATAGACTGTAAACATTTCTCAGCGATA	5467
QY	1838	AAATCCCTTTTCATTTCTAATGTAATCTAATTAACCTTATTATTAAATTCGCTCAT	1897
DB	5466	AAATCCCTTTTCATTTCTAATGTAATCTAATTAACCTTATTATTAAATTCGCTCAT	5407
QY	1898	AAATTAATCTTTTCTTATTAGCAAAATGCGCGATTAGCACACCCCTTTATTCGGTT	1957
DB	5406	AAATTAATCTTTTCTTATTAGCAAAATGCGCGATTAGCACACCCCTTTATTCGGTT	5347
QY	1958	AATGGCCATGACAGCCATGATAATTAATTAATCTAGGAGAGTAAATAATAGTAACC	2017
DB	5346	AATGGCCATGACAGCCATGATAATTAATTAATCTAGGAGAGTAAATAATAGTAACC	5287
QY	2018	AACATGATTAACAATATTAGAGTATCGTTTCAAAATGGTATGCGTTTGACACATCCA	2077
DB	5286	AACATGATTAACAATATTAGAGTATCGTTTCAAAATGGTATGCGTTTGACACATCCA	5227
QY	2078	CTATATATCCGTGCTGCTCTGCTCAATCCCAATCCCAATTCCTAGCGATT	2137
DB	5226	CTATATATCCGTGCTGCTCTGCTCAATCCCAATTCCTAGCGATT	5167
QY	2138	CCAGAAGTTTCTCAGAGTCGGAAGTTGACAGACATTAAGAACTGGCAGATGGTCAAT	2197
DB	5166	CCAGAAGTTTCTCAGAGTCGGAAGTTGACAGACATTAAGAACTGGCAGATGGTCAAT	5107
QY	2198	AACCTGAGGAGAGTCTGATTCGTTAACTGCTTCAAGTTAAGCAAGCGCTGCTGAT	2257
DB	5106	AACCTGAGGAGAGTCTGATTCGTTAACTGCTTCAAGTTAAGCAAGCGCTGCTGAT	5047
QY	2258	AACAGATGCGATGATGACAGCAATCAACATGGCACTGCCATTCCTACCTGTACAGTCA	2317
DB	5046	AACAGATGCGATGATGACAGCAATCAACATGGCACTGCCATTCCTACCTGTACAGTCA	4987
QY	2318	AGGATGGTAGAATCTGTCGGTCTGTCAGACAGTATTAAGCCATTCGCTGCATATT	2377
DB	4986	AGGATGGTAGAATCTGTCGGTCTGTCAGACAGTATTAAGCCATTCGCTGCATATT	4927
QY	2378	CAACAGCTCTTCTACGATAGGGCAGCAAAATCGCATCGTGAACGTTTGGGCTTCTACCG	2437
DB	4926	CAACAGCTCTTCTACGATAGGGCAGCAAAATCGCATCGTGAACGTTTGGGCTTCTACCG	4867
QY	2438	ATTAGAGTGTGATACACTTCTCTAGTATCCACCTGATCATTAATCGGCAAAATAG	2497
DB	4866	ATTAGAGTGTGATACACTTCTCTAGTATCCACCTGATCATTAATCGGCAAAATAG	4807

QY	2498	AGAAAAATTGACANTGTGTAGCGGCCAATCTGATTCACCTGAGATGCATAATCTAGTA	2557
DB	4806	AGAAAAATTGACANTGTGTAGCGGCCAATCTGATTCACCTGAGATGCATAATCTAGTA	4747
QY	2558	GAATCTCTTCGCTATCAAAATTCACCTTCACCTTCACCTTCACCGTTTGTCCATTCATGC	2617
DB	4746	GAATCTCTTCGCTATCAAAATTCACCTTCACCTTCACCTTCACCGTTTGTCCATTCATGC	4687
QY	2618	TGAATCTCTTCCTCTCTCTGTGTGACACACATCATCTCAATATCCGAATAGGCCCAT	2677
DB	4686	TGAATCTCTCTCTCTCTCTGTGTGACACACATCATCTCAATATCCGAATAGGCCCAT	4627
QY	2678	CAGTCTGACGACCAAGAGAGCCATAAACAACCAATAGCTTAAACATCATCCCATATTTAT	2737
DB	4626	CAGTCTGACGACCAAGAGAGCCATAAACAACCAATAGCTTAAACATCATCCCATATTTAT	4567
QY	2738	CCAATATTCGTTTCTTAAATTCATGAACAAATCTTCAATCTTCTCTCTAGTCATATTA	2797
DB	4566	CCAATATTCGTTTCTTAAATTCATGAACAAATCTTCAATCTTCTCTCTAGTCATATTA	4507
QY	2798	TTGGTCCATTCATCTTCTCCTTTCATCCCTTTTCCAGATAATTTAGATTTGCTTTCTAAATA	2857
DB	4506	TTGGTCCATTCATCTTCTCCTTTCATCCCTTTTCCAGATAATTTAGATTTGCTTTCTAAATA	4447
QY	2858	AGAAATATTTGGAGAGCACCGTTTCTTATTACGCTATTAAATACTCTTCTTCTAGCATCC	2917
DB	4446	AGAAATATTTGGAGAGCACCGTTTCTTATTACGCTATTAAATACTCTTCTTCTAGCATCC	4387
QY	2918	TTCAATTCCTTTTAAATAACAAATTTAGCATCTTAATCTTCAACAAATGCGCGTTTGTGA	2977
DB	4386	TTCAATTCCTTTTAAATAACAAATTTAGCATCTTAATCTTCAACAAATGCGCGTTTGTGA	4327
QY	2978	ACTACTCTTTTAAATAAATAATTTTCCGTTTCCCAATTCACATTCGCAATAATAGAAAATC	3037
DB	4326	ACTACTCTTTTAAATAAATAATTTTCCGTTTCCCAATTCACATTCGCAATAATAGAAAATC	4267
QY	3038	CATCTTCATCGGCTTTTTCGTCATCATCTGTATGAATCAAAATCGCTTCTCTGTGTAT	3097
DB	4266	CATCTTCATCGGCTTTTTCGTCATCATCTGTATGAATCAAAATCGCTTCTCTGTGTAT	4207
QY	3098	CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCACTAGGAGATTAACTTTTAC	3157
DB	4206	CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCACTAGGAGATTAACTTTTAC	4147
QY	3158	GGTGAACCTTCTCCAAATCAGACAAACCGTTTCAAAATCTTCTTCTCATCATCGGTCA	3217
DB	4146	GGTGAACCTTCTCCAAATCAGACAAACCGTTTCAAAATCTTCTTCTCATCATCGGTCA	4087
QY	3218	TAAATCCGATCTTTCAGGATATTTTTCAGATTTTTCGATTCGCGATTTGATATCGG	3277
DB	4086	TAAATCCGATCTTTCAGGATATTTTTCAGATTTTTCGATTCGCGATTTGATATCGG	4027
QY	3278	ATTATATTTATTTTTCGGTATTTTATTTTAAACGCTCTCA	3318
DB	4026	ATTATATTTATTTTTCGGTATTTTATTTTAAACGCTCTCA	3986

RESULT 3  
US-10-402-466A-1/c  
Sequence 1, Application US/10402466A  
Publication No. US20040028695A1  
GENERAL INFORMATION:  
APPLICANT: Park, Sukjoon  
APPLICANT: Giri, Lallan  
TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING  
AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS  
FILE REFERENCE: 18933-00005  
CURRENT APPLICATION NUMBER: US/10/402,466A  
CURRENT FILING DATE: 2003-03-20  
PRIOR APPLICATION NUMBER: 60/372,152  
PRIOR FILING DATE: 2002-04-12  
NUMBER OF SEQ ID NOS: 34

SEQ ID NO 1  
LENGTH: 6694  
TYPE: DNA  
ORGANISM: Artificial  
FEATURE:  
NAME/KEY:  
LOCATION:  
OTHER INFORMATION: Artificial DNA sequence to be used as one of the two backbones of  
OTHER INFORMATION: E. coli-Bacillus anthracis shuttle vectors. Designated as pBP I.  
US-10-402-466A-1

Query Match 29.0%; Score 1493; DB 13; Length 6694;  
Best Local Similarity 98.1%; Pred. No. 2.9e-282;  
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1778 TTCCTATCTTCATATAAGGTAACATATTGCGGGATAGACTGTAAACATCTCTACGCATA 1837  
DB 6355 TACTCTTTTAAATCCCGACTGGCAATGCGGGATAGACTGTAAACATCTCTACGCATA 6296

QY 1838 AAATCCCTTTTCATTTCTAATGTAAATCTATTACCTTATTATTAAATCAATTCGCTCAT 1897  
DB 6295 AAATCCCTTTTCATTTCTAATGTAAATCTATTACCTTATTATTAAATCAATTCGCTCAT 6236

QY 1898 AATTAATCCTTTTCTTATTAGGAAATGCCGATTTAAGCACACCCCTTTATTCGGTT 1957  
DB 6235 AATTAATCCTTTTCTTATTAGGAAATGCCGATTTAAGCACACCCCTTTATTCGGTT 6176

QY 1958 AATGCGCCATGACAGCCCATGATAATTAATACTACTAGGAGAGTTAATAATACGTAACC 2017  
DB 6175 AATGCGCCATGACAGCCCATGATAATTAATACTACTAGGAGAGTTAATAATACGTAACC 6116

QY 2018 AACATGATTAAACATTTAGAGTTCATGTTCAAAATGGTATGGTTTTCACACATCCA 2077  
DB 6115 AACATGATTAAACATTTAGAGTTCATGTTCAAAATGGTATGGTTTTCACACATCCA 6056

QY 2078 CTATATATCCGTTGCTGTTCTGTCCTACTCTGAATCCCATTCAGAAATTCCTAGCGATT 2137  
DB 6055 CTATATATCCGTTGCTGTTCTGTCCTACTCTGAATCCCATTCAGAAATTCCTAGCGATT 5996

QY 2138 CCAGAAATTTCTCAGAGTCGGAAGTTGACACAGATTAACGAACTGGACAGATGGTTCAT 2197  
DB 5995 CCAGAAATTTCTCAGAGTCGGAAGTTGACACAGATTAACGAACTGGACAGATGGTTCAT 5936

QY 2198 AACCTGAGGAGATCTGATCTGTTAACTGCTTCACTAGTTAAGACGAGCGCTCGTCTGAT 2257  
DB 5935 AACCTGAGGAGATCTGATCTGTTAACTGCTTCACTAGTTAAGACGAGCGCTCGTCTGAT 5876

QY 2258 AACAGATGCGATGATGACAGACCAATCAACATGSCACCTGCCATTCCTACCTGTACAGTCA 2317  
DB 5875 AACAGATGCGATGATGACAGACCAATCAACATGSCACCTGCCATTCCTACCTGTACAGTCA 5816

QY 2318 AGGATGGTAGAAATGTTGTCGTCCTTGACACAGATTAATAGCCATTTGCTGCTCATATT 2377  
DB 5815 AGGATGGTAGAAATGTTGTCGTCCTTGACACAGATTAATAGCCATTTGCTGCTCATATT 5756

QY 2378 CAAACAGCTCTTCTACGATAAGGCAATCGCATCGTGAACCTTTGGGCTTCTACCG 2437  
DB 5755 CAAACAGCTCTTCTACGATAAGGCAATCGCATCGTGAACCTTTGGGCTTCTACCG 5696

QY 2438 ATTTAGCAGTTGATACACTTTCTTAAGTATCCACCTGAATCATTAATTCGCAAAATAG 2497  
DB 5695 ATTTAGCAGTTGATACACTTTCTTAAGTATCCACCTGAATCATTAATTCGCAAAATAG 5636

QY 2498 AGAAAAATTTGACCATGTGAAGCGGCAATCTGATTTCCACCTGAGATGATTAATCTAGTA 2557  
DB 5635 AGAAAAATTTGACCATGTGAAGCGGCAATCTGATTTCCACCTGAGATGATTAATCTAGTA 5576

QY 2558 GAATCTCTTCGCTATCAAAATTTCACTTCCACCTTCCACTCCCGGTTGCTCAATTCATGCG 2617  
DB 5575 GAATCTCTTCGCTATCAAAATTTCACTTCCACCTTCCACTCCCGGTTGCTCAATTCATGCG 5516

QY 2618 TGAACCTCTGCTCTCTGTTGATGATGACACATCATCTCAATATCCGATAGGCGCCAT 2677

## RESULT 4

US-10-462-128-3/c  
Sequence 3, Application US/10462128  
Publication No. US20040067579A1

## GENERAL INFORMATION:

TITLE OF INVENTION: PROCESS FOR OXIDATION OF STEROIDS AND  
GENETICALLY ENGINEERED CELLS USED THEREIN

NUMBER OF SEQUENCES: 79

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIERMAN & MUSERLIAN

STREET: 600 THIRD AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/462,128

FILING DATE: 16-Jun-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/054,185

FILING DATE: 26-APR-1993  
APPLICATION NUMBER: US/08/002,608  
FILING DATE: 11-JAN-1993  
APPLICATION NUMBER: US/07/474,857  
FILING DATE: 30-OCT-1990  
APPLICATION NUMBER: US/07/474,798  
FILING DATE: 16-JULY-1990  
APPLICATION NUMBER: PCT/NL89/00072  
FILING DATE: 25-SEPT-1989  
APPLICATION NUMBER: NL/88/200904.6  
FILING DATE: 06-MAY-88  
APPLICATION NUMBER: NL/88/202080.3  
FILING DATE: 03-SEP-88  
ATTORNEY/AGENT INFORMATION:  
NAME: CHARLES A. MUSERLIAN  
REGISTRATION NUMBER: 19,683  
REFERENCE/DOCKET NUMBER: 146.1169 CON-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 661-8000  
TELEFAX: (212) 661-8002  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7336  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE  
TOPOLOGY: UNKNOWN  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: PLASMID pBHA-1  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-462-128-3

Query Match 29.0%; Score 1493; DB 13; Length 7336;  
Best Local Similarity 98.1%; Pred. No. 3.1e-262;  
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY	1778	TTCTTATCTGATATAAGGTAACTATTCGGGATAGACTGTAACATCTCCAGCATA	1837
DB	6367	TACTCTTTTAAATATCCCGACTGGCAATGCCGATAGACTGTAACATCTCCAGCATA	6308
QY	1838	AAATCCCTTTTCATTTTCTAATGTAATCTATTACCTTATTATTAAATTCGCTCAT	1897
DB	6307	AAATCCCTTTTCATTTTCTAATGTAATCTATTACCTTATTATTAAATTCGCTCAT	6248
QY	1898	AATTAATCGTTTCTTATTAGCAAAATGCGCGATTAAGCACACCTTTATTTCGTT	1957
DB	6247	AATTAATCGTTTCTTATTAGCAAAATGCGCGATTAAGCACACCTTTATTTCGTT	6188
QY	1958	AATGCGCCATGACAGCCATGATAATTAATACTAGGAGAGTAAATAAGTAACC	2017
DB	6187	AATGCGCCATGACAGCCATGATAATTAATACTAGGAGAGTAAATAAGTAACC	6128
QY	2018	AACATGATTAACAATATTAGAGTCATCGTTCAAAATGGTATGCGTTTTCACATCCA	2077
DB	6127	AACATGATTAACAATATTAGAGTCATCGTTCAAAATGGTATGCGTTTTCACATCCA	6068
QY	2078	CTATATATCCGTTGCTGTCTGTCACCTCTGTAATCCCATTCAGAAATTCCTAGCGATT	2137
DB	6067	CTATATATCCGTTGCTGTCTGTCACCTCTGTAATCCCATTCAGAAATTCCTAGCGATT	6008
QY	2138	CCAGAAGTTTCTCAGATCGGAAGTTGACAGACATTTACGAACTGGCAGAGTGGTCAT	2197
DB	6007	CCAGAAGTTTCTCAGATCGGAAGTTGACAGACATTTACGAACTGGCAGAGTGGTCAT	5948
QY	2198	AACCTGAAGGAGATCTGATTTGCTTAATCTGCTTACGTTTACGACCGAGCGCTCGCTGAT	2257
DB	5947	AACCTGAAGGAGATCTGATTTGCTTAATCTGCTTACGTTTACGACCGAGCGCTCGCTGAT	5888
QY	2258	AACAGATCGATGATGACAGACCAATCAACATGGCACCTGCCATTTGCTACCTGTACATCA	2317
DB	5887	AACAGATCGATGATGACAGACCAATCAACATGGCACCTGCCATTTGCTACCTGTACATCA	5828

RESULT 5

US-10-084-814-3/c

; Sequence 3, Application US/10084814

QY	2318	AGGATGGTAGAAATGTTTCGGTCTTCCACACGAAATATTAGGCATTTGCTCGCATATT	2377
DB	5827	AGGATGGTAGAAATGTTTCGGTCTTCCACACGAAATATTAGGCATTTGCTCGCATATT	5768
QY	2378	CAAAACAGCTCTTCTAGATAGGGCACAAATCGCATCTGTGAAAGTTGGGCTTCTACCG	2437
DB	5767	CAAAACAGCTCTTCTAGATAGGGCACAAATCGCATCTGTGAAAGTTGGGCTTCTACCG	5708
QY	2438	ATTAGCAGTTTGATACACATTTTCTCTAAGTATCCACCTTGAATCATATAATCGGAAAATAG	2497
DB	5707	ATTAGCAGTTTGATACACATTTTCTCTAAGTATCCACCTTGAATCATATAATCGGAAAATAG	5648
QY	2498	AGAAAAATTGACATGTGTAGCGGCCAATCTGATTCACCTGAGATGATTAATCTAGTA	2557
DB	5647	AGAAAAATTGACATGTGTAGCGGCCAATCTGATTCACCTGAGATGATTAATCTAGTA	5588
QY	2558	GAATCTCTTCGGTATCAAAATTCACCTTCCACCTTCCACTCCCGGTTGTCATTCAATGCC	2617
DB	5587	GAATCTCTTCGGTATCAAAATTCACCTTCCACCTTCCACTCCCGGTTGTCATTCAATGCC	5528
QY	2618	TGAATCTCTTCCTCTGTTGACATGACACATCATCTCAATATCCGAATAGGCCCCAT	2677
DB	5527	TGAATCTCTTCCTCTGTTGACATGACACATCATCTCAATATCCGAATAGGCCCCAT	5468
QY	2678	CAGTCTGACGACCAAGAGAGCCATAAACAACAATAGCCCTTAACATCATCCCATATTTAT	2737
DB	5467	CAGTCTGACGACCAAGAGAGCCATAAACAACAATAGCCCTTAACATCATCCCATATTTAT	5408
QY	2738	CCAAATATTCGTTCTTAAATTTTCATGAAACAAATCTTCAATCTTTCTCTAGTCATTATA	2797
DB	5407	CCAAATATTCGTTCTTAAATTTTCATGAAACAAATCTTCAATCTTTCTCTAGTCATTATA	5348
QY	2798	TTGTGTCATTTCACTATCTCTTCCCTTTTCAGATAAATTTTAGATTTTCTTTCTAAATA	2857
DB	5347	TTGTGTCATTTCACTATCTCTTCCCTTTTCAGATAAATTTTAGATTTTCTTTCTAAATA	5288
QY	2858	AGAATATTTGGAGAGACACCGTTTCTTTATTCAGCTATTATAAATCTGCTTCTCAAGCATCC	2917
DB	5287	AGAATATTTGGAGAGACACCGTTTCTTTATTCAGCTATTATAAATCTGCTTCTCAAGCATCC	5228
QY	2918	TTCAATCTTTTAAATAACAATATTAGCATCTTAATCTTCAACAAATGCGCGTTTGTGA	2977
DB	5227	TTCAATCTTTTAAATAACAATATTAGCATCTTAATCTTCAACAAATGCGCGTTTGTGA	5168
QY	2978	ACTACTCTTTTAAATAAATTTTTCGGTTCCCAATTCACATTCGAATAATAGAAAATC	3037
DB	5167	ACTACTCTTTTAAATAAATTTTTCGGTTCCCAATTCACATTCGAATAATAGAAAATC	5108
QY	3038	CATCTTCATCGGCTTTTCGTCATCATCTGTATGATCAAAATCGCCTTCTTCTGTGTAT	3097
DB	5107	CATCTTCATCGGCTTTTCGTCATCATCTGTATGATCAAAATCGCCTTCTTCTGTGTAT	5048
QY	3098	CAAGGTTTAAATTTTATGTAATTTCTTTTAAACAAACCAATAGGAGATTAAACCTTTTAC	3157
DB	5047	CAAGGTTTAAATTTTATGTAATTTCTTTTAAACAAACCAATAGGAGATTAAACCTTTTAC	4988
QY	3158	GGTGTAAACCTTTCCTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCA	3217
DB	4987	GGTGTAAACCTTTCCTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCA	4928
QY	3218	TAAATCCGATTCCTTTTACAGGATATTTTGCAGTTTTCGTAATTCGCGATTGTATATCGG	3277
DB	4927	TAAATCCGATTCCTTTTACAGGATATTTTGCAGTTTTCGTAATTCGCGATTGTATATCGG	4868
QY	3278	ATTATATTTATTTTTCGTAATTTTATTAATAACGTCATCA	3318
DB	4867	ATTATATTTATTTTTCGTAATTTTATTAATAACGTCATTTTACA	4827

```

/ Publication No. US2003010892A1
/
/ GENERAL INFORMATION:
/ APPLICANT: SLJJKHUIS, HERMAN; SELTEN,
/ GERARDUS CORNELIS MARIA; SMAAL,
/ ERIC BASTIAAN
/
/ TITLE OF INVENTION: PROCESS FOR OXIDATION OF
/ STEROIDS AND GENETICALLY ENGINEERED CELLS
/ USED THEREIN
/
/ NUMBER OF SEQUENCES: 79
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BIERMAN, MUSERLIAN & LUCAS
/ STREET: 600 THIRD AVENUE
/ CITY: NEW YORK
/ STATE: NEW YORK
/ COUNTRY: USA
/ ZIP: 10016
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: FLOPPY DISK
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: MICROSOFT WORD 97
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/084,814
/ FILING DATE: 26-Feb-2002
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/418,085
/ FILING DATE: 06-APR-1995
/ APPLICATION NUMBER: 08/054,185
/ FILING DATE: 26-APR-1993
/ APPLICATION NUMBER: 08/002,608
/ FILING DATE: 11-JAN-1993
/ APPLICATION NUMBER: 07/474,857
/ FILING DATE: 30-OCT-1990
/ APPLICATION NUMBER: 07/474,798
/ FILING DATE: 16-JULY-1990
/ APPLICATION NUMBER: FCI/NL89/00072
/ FILING DATE: 25-SEPT-1989
/ APPLICATION NUMBER: NL88/200904.6
/ FILING DATE: 06-MAY-1988
/ APPLICATION NUMBER: NL/88/202080.3
/ FILING DATE: 03-SEP-1988
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CHARLES A. MUSERLIAN
/ REGISTRATION NUMBER: 19,683
/ REFERENCE/DOCKET NUMBER: 146.1169-
/ CON-1-DIV-1
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 661-8000
/ TELEFAX: (212) 661-8002
/
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7336 BASE PAIRS
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: DOUBLE
/ TOPOLOGY: UNKNOWN
/
/ FEATURE:
/ OTHER INFORMATION: PLASMID pBHA-1
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
/
/ US-10-084-814-3
/
/ Query Match 29.0%; Score 1493; DB 15; Length 7336;
/ Best Local Similarity 98.1%; Pred.No.3.le-982;
/ Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
/
/ QY 1778 TTCTTATCTTGATAATAAGGGTAACCTATTCCCGGATAGACTGTAAACATTCTCACGCATA 1837
/ Db 6367 TACTCTTTTAAATATCCCGACTGGCAATGCCGGATAGACTGTAAACATTCTCACGCATA 6308
/
/ QY 1838 AAATCCCGTTTCATTTTCTAATGTAAATCTATTACCTTATTATTAATTCGAATTCGCTCAT 1897
/ Db 6307 AAATCCCGTTTCATTTTCTAATGTAAATCTATTACCTTATTATTAATTCGAATTCGCTCAT 6248
/
/ QY 1898 AATTAATCCCTTTTCTTTATTACGCAAAATGCCCGGATTTTAAGCACACACCCCTTTATTCGGTT 1957

```





QY 3218 TAAATCCGATCCTTTACAGGATATTTTCAGATTTTCCTCAATTCGCGATTGTATATCCG 3277  
DB 4927 TAAATCCGATCCTTTACAGGATATTTTCAGATTTTCCTCAATTCGCGATTGTATATCCG 4868  
QY 3278 ATTATATATTTATTTTCGGTATTTTATTTATTTAAACGCTCTCA 3318  
DB 4867 ATTATATATTTATTTTCGGTCAATCTTTGAACTTTTACA 4827  
RESULT 7  
US-10-402-466A-7/c  
; Sequence 7, Application US/10402466A  
; Publication No. US20040028695A1  
; GENERAL INFORMATION:  
; APPLICANT: Park, Sukjoon  
; APPLICANT: Giri, Lalilan  
; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING  
; TITLE OF INVENTION: AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS  
; FILE REFERENCE: 18933-00005  
; CURRENT APPLICATION NUMBER: US/10/402,466A  
; CURRENT FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: 60/372,152  
; PRIOR FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 34  
; SEQ ID NO 7  
; LENGTH: 8198  
; TYPE: DNA  
; ORGANISM: Bacillus anthracis  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; OTHER INFORMATION: Expression vector for B. anthracis Protective Antigen (PA). Ent  
; OTHER INFORMATION: sequence is shown since the vector sequence is different from the  
; OTHER INFORMATION: of the pBP vectors. The PA coding sequence is from 3735 to 5942.  
US-10-402-466A-7  
Query Match 29.0%; Score 1493; DB 13; Length 8198;  
Best Local Similarity 98.1%; Pred. No. 3.3e-282;  
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 1778 TTCTTATCTTGATAATAAGGTAACCTATTGCGGGATAGACTGTAAACATTTCTCAGGCATA 1837  
DB 3395 TACTCTTTTAAATATCCCGACTGGCAATGCGGGATAGACTGTAAACATTTCTCAGGCATA 3336  
QY 1838 AAATCCCTTTTCATTTCTTAATGTAAATCTATTTACCTTATTTATTTAAATCAATTCGCTCAT 1897  
DB 3335 AAATCCCTTTTCATTTCTTAATGTAAATCTATTTACCTTATTTATTTAAATCAATTCGCTCAT 3276  
QY 1898 AATTATCTCTTTTCTTATTACGCAAAATGCGCGATTTAGCACACCTTTATTCGGTT 1957  
DB 3275 AATTATCTCTTTTCTTATTACGCAAAATGCGCGATTTAGCACACCTTTATTCGGTT 3216  
QY 1958 AATGCGCCATGACAGCCATGATAATTAATACTACTAGGAGAAAGTTAAATAATACGTAACC 2017  
DB 3215 AATGCGCCATGACAGCCATGATAATTAATACTACTAGGAGAAAGTTAAATAATACGTAACC 3156  
QY 2018 AACATGATTAAACATTTATAGAGTTCATCGTTCAAAATGGTATGGTTTGCACATCCA 2077  
DB 3155 AACATGATTAAACATTTATAGAGTTCATCGTTCAAAATGGTATGGTTTGCACATCCA 3096  
QY 2078 CTATATATCGTCTGCTCTGCTCCACTCTCGAATCCATTCAGAAATTTCTTAGCGATT 2137  
DB 3095 CTATATATCGTCTGCTCTGCTCCACTCTCGAATCCATTCAGAAATTTCTTAGCGATT 3036  
QY 2138 CCAGAAAGTTTCTCAGAGTCGGAAGTTGACCGACATTTACGAATGGCCACAGATGTCAT 2197  
DB 3035 CCAGAAAGTTTCTCAGAGTCGGAAGTTGACCGACATTTACGAATGGCCACAGATGTCAT 2976  
QY 2198 AACCTGAAGCAAGATCTGATTGCTTAACTGCTTCACTTAAGACCGAAGCGCTCGTCGAT 2257  
DB 2975 AACCTGAAGCAAGATCTGATTGCTTAACTGCTTCACTTAAGACCGAAGCGCTCGTCGAT 2916  
QY 2258 AACAGATGCGATGATGACAGACCAATCAACATGSCACCTGCCATTTGCTACCTGTACAGTCA 2317

DB 2915 AACAGATGCGATGATGACAGACCAATCAACATGGCACCTGCCATTTCTACCTACGTACAGTCA 2856  
QY 2318 AGATGGTAGAATCTGTCGGTCTTCACACAGCAATATTTAGCCATTTGCTCGCATATT 2377  
DB 2855 AGATGGTAGAATCTGTCGGTCTTCACACAGCAATATTTAGCCATTTGCTCGCATATT 2796  
QY 2378 CAAACAGCTCTTTCTACGATAAGGGCACAAATCGCATCTGTGGAAACGTTTGGGCTTCTACCG 2437  
DB 2795 CAAACAGCTCTTTCTACGATAAGGGCACAAATCGCATCTGTGGAAACGTTTGGGCTTCTACCG 2736  
QY 2438 ATTAGCAGTTGATACACTTTCTTAAGTATCCACCTGAATCATATAATTCGCAAAATAG 2497  
DB 2735 ATTAGCAGTTGATACACTTTCTTAAGTATCCACCTGAATCATATAATTCGCAAAATAG 2676  
QY 2498 AGAAAATTTGACCATGTGTAAGCGGCCAATCTGATTTCCACCTGAGATGCATATCTAGTA 2557  
DB 2675 AGAAAATTTGACCAATGTGTAAGCGGCCAATCTGATTTCCACCTGAGATGCATATCTAGTA 2616  
QY 2558 GAATCTCTTGGCTATCAAAATTCATTTCCACCTTCCACTCAGCGTTGTCCATTCATGGC 2617  
DB 2615 GAATCTCTTGGCTATCAAAATTCATTTCCACCTTCCACTCAGCGTTGTCCATTCATGGC 2556  
QY 2618 TGAACCTCTGCTCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGGGCCAT 2677  
DB 2555 TGAACCTCTGCTCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGGGCCAT 2496  
QY 2678 CAGTCTGACGACCAAGAGGCCAATAAACACCAATAGCCTTAACATCATCCCCATATTTAT 2737  
DB 2495 CAGTCTGACGACCAAGAGGCCAATAAACACCAATAGCCTTAACATCATCCCCATATTTAT 2436  
QY 2738 CCAATATTCGTTCTTAATTTTCATGACCAATCTTCAATCTTCTTCTCTAGTCAATTTA 2797  
DB 2435 CCAATATTCGTTCTTAATTTTCATGACCAATCTTCAATCTTCTTCTCTAGTCAATTTA 2376  
QY 2798 TTGGTCCATTCACATTTCTCCCTTTTCAGATAAATTTAGATTTGCTTTTCTAAATA 2857  
DB 2375 TTGGTCCATTCACATTTCTCCCTTTTCAGATAAATTTAGATTTGCTTTTCTAAATA 2316  
QY 2858 AGAATTTGGAGAGACCGCTTCTTATTCAGCTATTAATAAATCTCTTCTTCTAGCATCC 2917  
DB 2315 AGAATTTGGAGAGACCGCTTCTTATTCAGCTATTAATAAATCTCTTCTTCTAGCATCC 2256  
QY 2918 TTCAATCTTTTAAATAAATTTTCGTTTCCCAATTCGATGATCAATCGCTTCTTCTGTGTCTAT 2977  
DB 2255 TTCAATCTTTTAAATAAATTTTCGTTTCCCAATTCGATGATCAATCGCTTCTTCTGTGTCTAT 2196  
QY 2978 ACTACTCTTTTAAATAAATTTTCGTTTCCCAATTCGATGATCAATCGCTTCTTCTGTGTCTAT 3037  
DB 2195 ACTACTCTTTTAAATAAATTTTCGTTTCCCAATTCGATGATCAATCGCTTCTTCTGTGTCTAT 2136  
QY 3038 CATCTTTCATCGGCTTTTTCGTCATCATCTGATGATCAATCGCTTCTTCTGTGTCTAT 3097  
DB 2135 CATCTTTCATCGGCTTTTTCGTCATCATCTGATGATCAATCGCTTCTTCTGTGTCTAT 2076  
QY 3098 CAAGTTTAAATTTTATGTATTTCTTTTAAACAAACCAATAGGAGATTAACTTTTAC 3157  
DB 2075 CAAGTTTAAATTTTATGTATTTCTTTTAAACAAACCAATAGGAGATTAACTTTTAC 2016  
QY 3158 GGTGTAAACCTTCTCCAAATTCAGACAAAGCTTTCAAAATTTCTTCTTCTCATCGGTCA 3217  
DB 2015 GGTGTAAACCTTCTCCAAATTCAGACAAAGCTTTCAAAATTTCTTCTTCTCATCGGTCA 1956  
QY 3218 TAAATCCGATCTCTTACAGGATATTTTCGAGTTTTCGTCATTTGCGGATTTGTATATCCG 3277  
DB 1955 TAAATCCGATCTCTTACAGGATATTTTCGAGTTTTCGTCATTTGCGGATTTGTATATCCG 1896  
QY 3278 ATTTATATTTATTTTTCGGTATTTTATTTTATTTAAACGCTCTCA 3318  
DB 1895 ATTTATATTTATTTTTCGGTCAATCATTTTGAACCTTTTACA 1855

RESULT 8



FILE REFERENCE: 9286.6CT  
CURRENT APPLICATION NUMBER: US/10/385,415  
CURRENT FILING DATE: 2003-03-10  
PRIOR APPLICATION NUMBER: US 09/936,028  
PRIOR FILING DATE: 2002-01-31  
PRIOR APPLICATION NUMBER: PCT/EP00/01899  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: DE 19910102.7  
PRIOR FILING DATE: 1999-03-08  
NUMBER OF SEQ ID NOS: 154  
SOFTWARE: Patent in version 3.2  
SEQ ID NO 2  
LENGTH: 5302  
TYPE: DNA  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: p6021-CAT Expression vector  
US-10-385-415-2  
Query Match 28.9%; Score 1487.6; DB 16; Length 5302;  
Best Local Similarity 99.1%; Pred. No. 3e-281;  
Matches 1496; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
QY 1809 CGGATAGACGTAAACATCTCAGCATAAATCCCTTCATTTCTTAATGTAATCTA 1868  
DB 2716 CGGATAGACGTAAACATCTCAGCATAAATCCCTTCATTTCTTAATGTAATCTA 2775  
QY 1869 TTACCTATTATTAATTCGCTCATATTAATTCCTTTTCTTATTAGCGCAAAATGG 1928  
DB 2776 TTACCTATTATTAATTCGCTCATATTAATTCCTTTTCTTATTAGCGCAAAATGG 2835  
QY 1929 CCGATTTAAGCACACCCCTTTATTCGGTTAATGGCCATGACAGCCATGATTAATCTAA 1988  
DB 2836 CCGATTTAAGCACACCCCTTTATTCGGTTAATGGCCATGACAGCCATGATTAATCTAA 2895  
QY 1989 TACTAGGAGAGTTAATAAATACGTAAACCAATGATTAACAAATATTAGAGGTCACTCGT 2048  
DB 2896 TACTAGGAGAGTTAATAAATACGTAAACCAATGATTAACAAATATTAGAGGTCACTCGT 2955  
QY 2049 TCAAAATGGTATGGCTTTTGACACATCATATATATCCGTGTCGTTCTGTCCTACTCTG 2108  
DB 2956 TCAAAATGGTATGGCTTTTGACACATCATATATATCCGTGTCGTTCTGTCCTACTCTG 3015  
QY 2109 ATCCCATTCAGAAATCTCTAGCGATTCCAGAGTTCTCAGAGTCGGAAGTTGACC 2168  
DB 3016 ATCCCATTCAGAAATCTCTAGCGATTCCAGAGTTCTCAGAGTCGGAAGTTGACC 3075  
QY 2169 AGACATTACGAATCGGCACAGATGTCATTAACCTGAAGGAATCTGATTCTTAACATGC 2228  
DB 3076 AGACATTACGAATCGGCACAGATGTCATTAACCTGAAGGAATCTGATTCTTAACATGC 3135  
QY 2229 TTCAGTTAAGCCGAAGCGCTCGTATACAGATGCGATGATGACAGCAATCAACAT 2288  
DB 3136 TTCAGTTAAGCCGAAGCGCTCGTATACAGATGCGATGATGACAGCAATCAACAT 3195  
QY 2289 GGCACCTGCATTTGCTTACCTGATACAGTCAAGGATGGTAGAATTTGTCGGTCTTGAC 2348  
DB 3196 GGCACCTGCATTTGCTTACCTGATACAGTCAAGGATGGTAGAATTTGTCGGTCTTGAC 3255  
QY 2349 AGCAATATTACGCATTTGCTTGCATTTCAAAACAGCTTTCTACGATAAGGGCAAAAT 2408  
DB 3256 AGCAATATTACGCATTTGCTTGCATTTCAAAACAGCTTTCTACGATAAGGGCAAAAT 3315  
QY 2409 CGCATCGTGAAGCTTTGGCTTCTACCGATTTAGCAGTTTCATACACTTTCTTAAGTA 2468  
DB 3316 CGCATCGTGAAGCTTTGGCTTCTACCGATTTAGCAGTTTCATACACTTTCTTAAGTA 3375  
QY 2469 TCCACCTGAATCATAAATCGCAAAATAGAGAAAATGACCATGTGTAAGCGGCAATC 2528  
DB 3376 TCCACCTGAATCATAAATCGCAAAATAGAGAAAATGACCATGTGTAAGCGGCAATC 3435  
QY 2529 TGAATCCACTGAGATGCAATATCTAGTAAATCTCTTCGCTATCAAAATTCACCTCCAC 2588

DB 3436 TGAATCCACTGAGATGCAATATCTAGTAGAATCTCTCGCTATCAAAATTCACCTCCAC 3495  
QY 2589 CTTCCACTCACCGGTTGTCCATTTCAATGGCTGAATCTCTCTCTCTGTGTGACATGACACA 2648  
DB 3496 CTTCCACTCACCGGTTGTCCATTTCAATGGCTGAATCTCTCTCTCTGTGTGACATGACACA 3555  
QY 2649 CATCATCTCAATATCCGAATAGGGCCCATCAGTCTGACGACCAAGAGAGCCATAAACACC 2708  
DB 3556 CATCATCTCAATATCCGAATAGGGCCCATCAGTCTGACGACCAAGAGAGCCATAAACACC 3615  
QY 2709 AATAGCCTTAACATCATCCCCATATTTATCCAAATATTCGTTCCCTTAATTTCAAGCAAT 2768  
DB 3616 AATAGCCTTAACATCATCCCCATATTTATCCAAATATTCGTTCCCTTAATTTCAAGCAAT 3675  
QY 2769 CTTCAATCT 2828  
DB 3676 CTTCAATCT 3735  
QY 2829 AGATAATTTTACATTTGCTTTTCTAAATAAGAAATATTTGGAGAGCACCGTCTCTTATTTCAG 2888  
DB 3736 AGATAATTTTACATTTGCTTTTCTAAATAAGAAATATTTGGAGAGCACCGTCTCTTATTTCAG 3795  
QY 2889 CTATTAAATAACTCGTCTCTCTTAAGCATCTTCAATCTCTTTTAAATAACATTTATAGCATCT 2948  
DB 3796 CTATTAAATAACTCGTCTCTCTTAAGCATCTTCAATCTCTTTTAAATAACATTTATAGCATCT 3855  
QY 2949 AATCTTCAACAAACTGGCCCGTTTGTGAACTACTCTTTAAATAAATAATTTTTCGTTTC 3008  
DB 3856 AATCTTCAACAAACTGGCCCGTTTGTGAACTACTCTTTAAATAAATAATTTTTCGTTTC 3915  
QY 3009 CCAATTTCCCAATGCAATAATAAGAAATCCATCTTTCATCGGCTTTTTCGTCATCATCTGT 3068  
DB 3916 CCAATTTCCCAATGCAATAATAAGAAATCCATCTTTCATCGGCTTTTTCGTCATCATCTGT 3975  
QY 3069 ATGAATCAAAATCGCTCTCTCTGTGTCATCAGGTTTAAATTTTATGTTATTTCTTTTAA 3128  
DB 3976 ATGAATCAAAATCGCTCTCTCTGTGTCATCAGGTTTAAATTTTATGTTATTTCTTTTAA 4035  
QY 3129 CAAACACCATAGAGATTAACCTTTTACGGTGTAAACCTTCTCCCAAAATCAGACAAACG 3188  
DB 4036 CAAACACCATAGAGATTAACCTTTTACGGTGTAAACCTTCTCCCAAAATCAGACAAACG 4095  
QY 3189 TTTCAAATCTTTTCTTCTCATFCATCGTCTATAAAATCCGATCCCTTTACAGGATATTTTGC 3248  
DB 4096 TTTCAAATCTTTTCTTCTCATFCATCGTCTATAAAATCCGATCCCTTTACAGGATATTTTGC 4155  
QY 3249 AGTTTCGTCGAATTCGGATTTATATCCGATTTATTTATTTTTCGGTATTTTATTATTA 3308  
DB 4156 AGTTTCGTCGAATTCGGATTTATATCCGATTTATTTATTTTTCGGTATTTTATTATTA 4215  
QY 3309 AAACGCTCTCA 3318  
DB 4216 AACTTTTACA 4225

RESULT 10  
US-10-385-415-4  
; Sequence 4, Application US/10385415  
; Publication No. US20040014158A1  
; GENERAL INFORMATION:  
; APPLICANT: Bacher, Adelbert  
; APPLICANT: Fischer, Markus  
; TITLE OF INVENTION: PROTEIN CONJUGATES, METHOD, VECTORS, PROTEINS AND DNA FOR  
; TITLE OF INVENTION: PRODUCING THEM, THEIR USE AND MEDICAMENTS AND VACCINES CONTAINING  
; TITLE OF INVENTION: A CERTAIN QUANTITY OF SAID PROTEIN CONJUGATES  
; FILE REFERENCE: 9286.6CT  
; CURRENT APPLICATION NUMBER: US/10/385,415  
; CURRENT FILING DATE: 2003-03-10  
; PRIOR APPLICATION NUMBER: US 09/936,028  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: PCT/EP00/01899  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: DE 19910102.7

```
; PRIOR FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 5767
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p602-BS-lusY Expression plasmid
US-10-385-415-4

Query Match      28.9%; Score 1487.6; DB 16; Length 5767;
Best Local Similarity 99.1%; Pred. No. 3.1e-281;
Matches 1496; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1809 CGGATAGACGTACATCTTCACGCAATAAAATCCCTTCATTTCTAATGTAAATCTA 1868
DB 3191 CGGATAGACGTACATCTTCACGCAATAAAATCCCTTCATTTCTAATGTAAATCTA 3240
QY 1869 TTACCTTATTATTAAATCAATTCCTCATATAATCAATCTTTTCTTATTACGCAAAATGG 1928
DB 3241 TTACCTTATTATTAAATCAATTCCTCATATAATCAATCTTTTCTTATTACGCAAAATGG 3300
QY 1929 CCCGATTTAAGCAGACACCTTTTATTCGCTTAATGCGCCATGACGCGCATGATTAATCTAA 1988
DB 3301 CCCGATTTAAGCAGACACCTTTTATTCGCTTAATGCGCCATGACGCGCATGATTAATCTAA 3360
QY 1989 TACTAGGAGAGATTAATAAATACGTAAACCAATGATTAACAATTTATAGAGGTTCATCGT 2048
DB 3361 TACTAGGAGAGATTAATAAATACGTAAACCAATGATTAACAATTTATAGAGGTTCATCGT 3420
QY 2049 TCAAAATGGTATGGTTTGTGACATATCCATATATATCCGTGCTGTTCTGTCCACTCTTG 2108
DB 3421 TCAAAATGGTATGGTTTGTGACATATCCATATATATCCGTGCTGTTCTGTCCACTCTTG 3480
QY 2109 AATCCCATTCAGAAATTCCTTAGCGATTCCAGAAATTCCTAGAGTCGGAAGTTGAC 2168
DB 3481 AATCCCATTCAGAAATTCCTTAGCGATTCCAGAAATTCCTAGAGTCGGAAGTTGAC 3540
QY 2169 AGAATTCAGAACTGGCAGACAGATGGTCATAAATCTGAAGGAAGATCTGATTCCTTAATGTC 2228
DB 3541 AGAATTCAGAACTGGCAGACAGATGGTCATAAATCTGAAGGAAGATCTGATTCCTTAATGTC 3600
QY 2229 TTCAGTTAAGACCGAAGCGCTTCGTATATACAGATGCGATGATGACAGACCAATCAACAT 2288
DB 3601 TTCAGTTAAGACCGAAGCGCTTCGTATATACAGATGCGATGATGACAGACCAATCAACAT 3660
QY 2289 GGCACCTGCCATTCCTACCTGACAGTCAAGGATGGTAGAAATGTTCTCGGTCTTGCAC 2348
DB 3661 GGCACCTGCCATTCCTACCTGACAGTCAAGGATGGTAGAAATGTTCTCGGTCTTGCAC 3720
QY 2349 AGCAATATTAGCCATTTGCTGATATCAACAGCTCTTCTACGATAGAGGACCAAT 2408
DB 3721 AGCAATATTAGCCATTTGCTGATATCAACAGCTCTTCTACGATAGAGGACCAAT 3780
QY 2409 CGCATCGTGGAACTGTTGGGCTTCTACCGATTAGCAGATTGTGATACACTTTCTCTAAGTA 2468
DB 3781 CGCATCGTGGAACTGTTGGGCTTCTACCGATTAGCAGATTGTGATACACTTTCTCTAAGTA 3840
QY 2469 TCCACCTGAATCATAAATTCGGCAAAATAGAGAAAATGACCATGTGTAGCGGCCAATC 2528
DB 3841 TCCACCTGAATCATAAATTCGGCAAAATAGAGAAAATGACCATGTGTAGCGGCCAATC 3900
QY 2529 TGATTTCCACCTGAGATGATTAATCTAGTAGAATCTCTTCGCTATCAAAATTCACCTCCAC 2588
DB 3901 TGATTTCCACCTGAGATGATTAATCTAGTAGAATCTCTTCGCTATCAAAATTCACCTCCAC 3960
QY 2589 CTTCCACTCAGCGGTGTCCATATCATGGCTGAATCTGCTCTCTCTGTTGACATGACACA 2648
DB 3961 CTTCCACTCAGCGGTGTCCATATCATGGCTGAATCTGCTCTCTCTGTTGACATGACACA 4020
QY 2649 CATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATAAACAC 2708
```

```
4021 CATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATAAACACC 4080
QY 2709 AATAGCCTTAACATCATCCCATATTTTATCAATATTCGTTTCTTAATTTTCATGAACAAT 2768
DB 4081 AATAGCCTTAACATCATCCCATATTTTATCAATATTCGTTTCTTAATTTTCATGAACAAT 4140
QY 2769 CTTCAATCTTTCTCTCTAGTCATTTATTTGGTCCATCTACTATTCATCTCCCTTTTC 2828
DB 4141 CTTCAATCTTTCTCTCTAGTCATTTATTTGGTCCATCTACTATTCATCTCCCTTTTC 4200
QY 2829 AGATAATTTTAGATTTGCTTTCTTAATAAGAAATATTTGGAGAGACACCGTCTCTATTACAG 2888
DB 4201 AGATAATTTTAGATTTGCTTTCTTAATAAGAAATATTTGGAGAGACACCGTCTCTATTACAG 4260
QY 2889 CTATTAAATAACCTCGTCTTCTTAAGCATCTTCAATCTTTTAAATAACAATTTATAGCATCT 2948
DB 4261 CTATTAAATAACCTCGTCTTCTTAAGCATCTTCAATCTTTTAAATAACAATTTATAGCATCT 4320
QY 2949 AATCTTCAACAACTGGCGCGTGTGTAACACTCTTTTAATAAATAAATTTTCCGTTTC 3008
DB 4321 AATCTTCAACAACTGGCGCGTGTGTAACACTCTTTTAATAAATAAATTTTCCGTTTC 4380
QY 3009 CCAATTCACATGCAATAATAAGAAAATCCATCTTTCATCGGCTTTTTCGTCATCATCTGT 3068
DB 4381 CCAATTCACATGCAATAATAAGAAAATCCATCTTTCATCGGCTTTTTCGTCATCATCTGT 4440
QY 3069 ATGAATCAAAATCGCTTCTCTGTGTCATCAAGGTTAAATTTTATGATTTCTTTTAA 3128
DB 4441 ATGAATCAAAATCGCTTCTCTGTGTCATCAAGGTTAAATTTTATGATTTCTTTTAA 4500
QY 3129 CAAACACCATAGGAGATTAACCTTTTACGGTGTAAACCTTCTCCCAATCAGACAAACG 3188
DB 4501 CAAACACCATAGGAGATTAACCTTTTACGGTGTAAACCTTCTCCCAATCAGACAAACG 4560
QY 3189 TTTCAAAATCTTTTCTCATCATCGGTCAATAAAATCCGATATCTTTTTCGTTATTTTATTA 3248
DB 4561 TTTCAAAATCTTTTCTCATCATCGGTCAATAAAATCCGATATCTTTTTCGTTATTTTATTA 4620
QY 3249 AGTTTCGTCATTCGCGATGTTATATCGATTATATTTTTCGTTATTTTATTA 3308
DB 4621 AGTTTCGTCATTCGCGATGTTATATCGATTATATTTTTCGTTATTTTATTA 4680
QY 3309 AAAACGTCTCA 3318
DB 4681 AACTTTTACA 4690

RESULT 11
US-09-869-855A-3/c
; Sequence 3, Application US/09869855A
; Publication No. US20030044940A1
; GENERAL INFORMATION:
; APPLICANT: Rasmussen, Michael, Dolberg
; TITLE OF INVENTION: Method for Increasing Gene Copy Number
; FILE REFERENCE: 10028-204-US
; CURRENT APPLICATION NUMBER: US/09/869,855A
; CURRENT FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5793
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-869-855A-3

Query Match      23.5%; Score 1207.6; DB 10; Length 5793;
Best Local Similarity 98.5%; Pred. No. 2.8e-226;
Matches 1219; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1813 ATAGACTGTAACTCTCACGCATAAAATCCCTTTTCATTTCTTAATGTAATCTATTAC 1872
```

2069	AGAGATCTCAGGGTTGGCGCGCGCAAAA	TCCCTTTTCATTTCTTAATGTAAATCTATTAC	2010
1873	CTTATTATTAATCAATTCGCTCATAAATTAAT	CCTTTTTCCTTATTACGCAAAATGCGCCG	1932
2009	CTTATTATTAATCAATTCGCTCATAAATTAAT	CCTTTTTCCTTATTAGCGAAAATGCGCCG	1950
1933	ATTATAGCACACCCCTTTATTCGGTTAATCGCCAT	GACAGCCATGATAATTACTATACT	1992
1949	ATTTAAGCACACCCCTTTATTCGGTTAATCGCCAT	GACAGCCATGATAATTACTATACT	1890
1993	AGGAGAGTTTAATAAATPACGTAAACCAACATGAT	TAAACAAATTATTAGAGGTCATCGTTCAA	2052
1889	AGGAGAGTTTAATAAATPACGTAAACCAACATGAT	TAAACAAATTATTAGAGGTCATCGTTCAA	1830
2053	AATGGTATCGGTTTGGACACATCCACTATATAT	CCGTGTCGGTCTGCTCCACTCCTGAATC	2112
1829	AATGGTATCGGTTTGGACACATCCACTATATAT	CCGTGTCGGTCTGCTCCACTCCTGAATC	1770
2113	CCATTCCAGAAAATTCCTAGCGATTCAGAAAGTTT	CTCAGAGTCGGAAGTGTACACAGAC	2172
1769	CCATTCCAGAAAATTCCTAGCGATTCAGAAAGTTT	CTCAGAGTCGGAAGTGTACACAGAC	1710
2173	ATTACGAACHTGGCACAGATGGTCATACCTTGAA	AGGAAGATCTGATTTGCTTTAACTGCTTCA	2232
1709	ATTACGAACHTGGCACAGATGGTCATACCTTGAA	AGGAAGATCTGATTTGCTTTAACTGCTTCA	1650
2233	GTTAAGACCGAAGCGCTCGCTGATATAACAGAT	CGGATGTGAGCAAAATCAACATGGCA	2292
1649	GTTAAGACCGAAGCGCTCGCTGATATAACAGAT	CGGATGTGAGCAAAATCAACATGGCA	1590
2293	CCTGCCATTGCTACCTGTACAGTCAAGGATGGT	AGAAAATGTTGCGTCTCTTGCACAACGA	2352
1589	CCTGCCATTGCTACCTGTACAGTCAAGGATGGT	AGAAAATGTTGCGTCTCTTGCACAACGA	1530
2353	ATATTACGCCATTGCGCTGCATATTCAAACAGCT	CTTCTACAGTAAAGGACACAAATCGCA	2412
1529	ATATTACGCCATTGCGCTGCATATTCAAACAGCT	CTTCTACAGTAAAGGACACAAATCGCA	1470
2413	TCGTGGAACGTTTGGGCTTCTACCGATTTAGCAG	TTTGATACACTTTTCTCTAAGTATCCA	2472
1469	TCGTGGAACGTTTGGGCTTCTACCGATTTAGCAG	TTTGATACACTTTTCTCTAAGTATCCA	1410
2473	CCTGTAATCAATAATCGGCAAAATAGAGAAAAT	TGACCAATGTGTAAGCGGCCAATCTGAT	2532
1409	CCTGAATCAATAATCGGCAAAATAGAGAAAAT	TGACCAATGTGTAAGCGGCCAATCTGAT	1350
2533	TCACACCTGAGATGATAAATCTAGTAGAAATCT	CTTTGCGTATCAAAATTCACCTCCACCTTC	2592
1349	TCACACCTGAGATGATAAATCTAGTAGAAATCT	CTTTGCGTATCAAAATTCACCTCCACCTTC	1290
2593	CACCTACCCGTTGTCATTATGCGCTGAACTCTCT	TCCTCTGTGATGACACACATC	2652
1289	CACCTACCCGTTGTCATTATGCGCTGAACTCTCT	TCCTCTGTGATGACACACATC	1230
2653	ATCTCAATATCCGAATAGGCGCCATCAGCTGACG	CCAGAGAGCCATAAAGACCAATA	2712
1229	ATCTCAATATCCGAATAGGCGCCATCAGCTGACG	CCAGAGAGCCATAAAGACCAATA	1170
2713	GCCTTAAACATCATCCCATAATTTATCCAATATT	CGTTTTCCTTAAATTCATGAACAATCTTC	2772
1169	GCCTTAAACATCATCCCATAATTTATCCAATATT	CGTTTTCCTTAAATTCATGAACAATCTTC	1110
2773	ATTCCTTTCTCTCTAGTCATTATTATTGGTCCAT	TCTCACTTCTCCTTTTCAGAT	2832
1109	ATTCCTTTCTCTCTAGTCATTATTATTGGTCCAT	TCTCACTTCTCCTTTTCAGAT	1050
2833	AAATTTAGATTGCTTTTCTAAATAAGAAATATT	GGAGAGCACCGTTCTTATTACGCTAT	2892
1049	AAATTTAGATTGCTTTTCTAAATAAGAAATATT	GGAGAGCACCGTTCTTATTACGCTAT	990
2893	TAATAACTCGCTTCTCCTAAGCATCCTTCAAT	CCCTTTTAAATAACAATTTATAGCATCTAATC	2952
989	TAATAACTCGCTTCTCCTAAGCATCCTTCAAT	CCCTTTTAAATAACAATTTATAGCATCTAATC	930

```

Qy      2953  TTCAACAGAACTGCCCGCTTGTGTGAAGTACTCTCTTAATAAAAAATAATTTTCCCGTTCCCAA 3012
Db      929  TTCAACAGAACTGCCCGCTTGTGTGAAGTACTCTCTTAATAAAAAATAATTTTCCCGTTCCCAA 870

Qy      3013  TTCCACATTGCAATAATAGAAAAATCCCATCTTCATCGGC 3050
Db      869  TTCCACATTGCAATAATAGAAAAATCCCATCTTCATCGTC 832

RESULT 12
US-09-869-855A-2/c
; Sequence 2, Application US/09869855A
; Publication No. US20030044940A1
; GENERAL INFORMATION:
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Method For Increasing Gene Copy Number
; FILE REFERENCE: 10028.204-US
; CURRENT APPLICATION NUMBER: US/09/869,855A
; CURRENT FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 5943
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-869-855A-2

```

```
Query Match      23.5%; Score 1207.6; DB 10; Length 5943;
Best Local Similarity 98.5%; Pred. No. 2.8e-226;
Matches 1219; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

Qy	1813	ATAGACGTGAACATCTCAGCGATAAAATCCCTTTCATTTCTTAATGTAAATCTATTAC	1872
Db	2069	AGAGATCTCAGGTTGCGGCCCAAAATCCCTTTCACTTTCTAATGTAAATCTATTAC	2010
Qy	1873	CTTATTATTAATTCAATTCGCTCATTAATTAATCCTTTTCTTATTACGCAAAATGGCCCG	1932
Db	2009	CTTATTATTAATTCAATTCGCTCATTAATTAATCCTTTTCTTATTACGCAAAATGGCCCG	1950
Qy	1933	ATTTAAGCACACCTTTTATTCGGTTAATCGGCCATGACAGCCCATGATAATTAATAATPACT	1992
Db	1949	ATTTAAGCACACCTTTTATTCGGTTAATCGGCCATGACAGCCCATGATAATTAATAATPACT	1890
Qy	1993	AGGAGAAGTTAAATAATGCTAACCAACATGATTAACAATATTATAGAGTCACTGTTCAA	2052
Db	1889	AGGAGAAGTTAAATAAATPACGTAACCAACATGATTAAACAATATTATAGAGTCACTGTTCAA	1830
Qy	2053	AATGGTATGGTTTGTGACACATCCCATATATATCCGTTGTCGTTCTGTCACACTCCTGAATC	2112
Db	1829	AATGGTATGGTTTGTGACATCCCATATATATCCGTTGTCGTTCTGTCACACTCCTGAATC	1770
Qy	2113	CCATTCAGAGAAATTCCTACGCGATTCCAGAAAGTTTCTCAGAGTCGGAAGTTGACCAAGAC	2172
Db	1769	CCATTCAGAGAAATTCCTACGCGATTCCAGAAAGTTTCTCAGAGTCGGAAGTTGACCAAGAC	1710
Qy	2173	ATTACGAATGCGACAGATGGTGTAAATACCTGAAGGAAGATCTGATTCGTTAAATGCTTCA	2232
Db	1709	ATTACGAATGCGACAGATGGTGTAAATACCTGAAGGAAGATCTGATTCGTTAAATGCTTCA	1650
Qy	2233	GTTTAAGACCGAAGCGCTGTCGTATTAACAGATCGATGATGCAGACCAATCAACATGGCA	2292
Db	1649	GTTTAAGACCGAAGCGCTGTCGTATTAACAGATCGATGATGCAGACCAATCAACATGGCA	1590
Qy	2293	CCTGCCATTGCTACCTGTACAGTCAAGGATGTAGAAATGTTGTCGGTCCCTTGCACACGA	2352
Db	1589	CCTGCCATTGCTACCTGTACAGTCAAGGATGTAGAAATGTTGTCGGTCCCTTGCACACGA	1530
Qy	2353	ATATTACGCCATTTGCTGCGATATTCAACACAGCTCTTCTACGATAAGGGCACAAAATCGCA	2412
Db	1529	ATATTACGCCATTTGCTGCGATATTCAACACAGCTCTTCTACGATAAGGGCACAAAATCGCA	1470

2413 TCGTGGACGTTTGGGCTTCTACCGATTTAGCAGTTTGATACACTTCTCTAAGTATCCA 2472  
Db |||||||  
1469 TCGTGGACGTTTGGGCTTCTACCGATTTAGCAGTTTGATACACTTCTCTAAGTATCCA 1410  
Qy |||||||  
2473 CTTGAATCATAAATCGGCAAAATAGAGAAAATGACCAATGTGTGTAAGCGGCCAATCTGAT 2532  
Db |||||||  
1409 CTTGAATCATAAATCGGCAAAATAGAGAAAATGACCAATGTGTGTAAGCGGCCAATCTGAT 1350  
Qy |||||||  
2533 TCCACCTGAGATGATATCTAGTAGAATCTCTTCGGTATCAAAATCACTTCCACCTTC 2592  
Db |||||||  
1349 TCCACCTGAGATGATATCTAGTAGAATCTCTTCGGTATCAAAATCACTTCCACCTTC 1290  
Qy |||||||  
2593 CACTCACCGGTTGTCATTCTGATGCTGAACCTCTGCTTCTCTGTTGATGACGACACATC 2652  
Db |||||||  
1289 CACTCACCGGTTGTCATTCTGATGCTGAACCTCTGCTTCTCTGTTGATGACGACATC 1230  
Qy |||||||  
2653 ATCTCAATATCCGAATAGGCGCCATCTAGTCTGACGACCAAGAGAGCCATTAACCAATA 2712  
Db |||||||  
1229 ATCTCAATATCCGAATAGGCGCCATCTAGTCTGACGACCAAGAGAGCCATTAACCAATA 1170  
Qy |||||||  
2713 GCCTTAACATATCCCAATATTTATCCAAATATTCGTTCTTAAATTCATGAAACAATCTTC 2772  
Db |||||||  
1169 GCCTTAACATATCCCAATATTTATCCAAATATTCGTTCTTAAATTCATGAAACAATCTTC 1110  
Qy |||||||  
2773 ATCTTTCTCTCTAGTCAATATTTATTTGGTCCATCTCACTATCTCATCTCCCTTTTCAGAT 2832  
Db |||||||  
1109 ATCTTTCTCTCTAGTCAATATTTATTTGGTCCATCTCACTATCTCATCTCCCTTTTCAGAT 1050  
Qy |||||||  
2833 AATTTAGATTTGCTTTCTTAATAAGAAATATTTGGAGAGACCGTTCTTATTCAGCTAT 2892  
Db |||||||  
1049 AATTTAGATTTGCTTTCTTAATAAGAAATATTTGGAGAGACCGTTCTTATTCAGCTAT 990  
Qy |||||||  
2893 TAATAACTCTGTTCTTCTAAGCATCTTCAATCTCTTTTAATAACCAATATATAGCATCTAATC 2952  
Db |||||||  
989 TAATAACTCTGTTCTTCTAAGCATCTTCAATCTCTTTTAATAACCAATATATAGCATCTAATC 930  
Qy |||||||  
2953 TTCAACAACTGGCCGTTTGTGTAAGTACTCTTAAATAAATATTTTCCGTTCCCAA 3012  
Db |||||||  
929 TTCAACAACTGGCCGTTTGTGTAAGTACTCTTAAATAAATATTTTCCGTTCCCAA 870  
Qy |||||||  
3013 TTCCCATTTGCAATTAATAAGAAAATCCATCTTTCATCGGC 3050  
Db |||||||  
869 TTCCCATTTGCAATTAATAAGAAAATCCATCTTTCATCGTC 832

## RESULT 13

US-10-032-393-16  
; Sequence 16, Application US/10032393  
; Publication No. US20030027286A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Wall, Daniel  
; APPLICANT: Gross, Molly  
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
; FILE REFERENCE: ELITRA.010A  
; CURRENT APPLICATION NUMBER: US/10/032,393  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/259,434  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: 09/948,993  
; PRIOR FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 6852  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Vector pXyl-T5 p15a

US-10-032-393-16

Query Match 19.1%; Score 981.2; DB 15; Length 6852;  
Best Local Similarity 99.2%; Pred. No. 8.1e-182;  
Matches 986; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 831 TTAGAAAACCGACTGTCTTAAAGATGACGTCGGCATTAICTCATATATATAAAGCCAGTCAT 890  
Db |||||||  
4543 TTAGAAAACCGACTGTCTTAAAGATGACGTCGGCATTAICTCATATATATAAAGCCAGTCAT 4602  
Qy 891 TAGGCCTTATCTGACAAATTCCTGATAGAGTTTCATAAACAATCCCTGCATGATAAACCATCAC 950  
Db |||||||  
4603 TAGGCCTTATCTGACAAATTCCTGATAGAGTTTCATAAACAATCCCTGCATGATAAACCATCAC 4662  
Qy 951 AAACAGAAATGATGTACCTGTAAAGATAGCGGTAAATATATTAATTAACCTTTATTAATGA 1010  
Db |||||||  
4663 AAACAGAAATGATGTACCTGTAAAGATAGCGGTAAATATATTAATTAACCTTTATTAATGA 4722  
Qy 1011 ATTTTCCCTGCTAATAATGCTGAAGGTAAATTAATTAATTAATTAATTAATTAATTAATTA 1070  
Db |||||||  
4723 ATTTTCCCTGCTAATAATGCTGAAGGTAAATTAATTAATTAATTAATTAATTAATTAATTA 4782  
Qy 1071 ACCCAGTAATGAACTCCATGGAATATAGAAAGAGAAAAGCATTTTCAGGTATAGGTG 1130  
Db |||||||  
4783 ACCCAGTAATGAACTCCATGGAATATAGAAAGAGAAAAGCATTTTCAGGTATAGGTG 4842  
Qy 1131 TTTTGGGAAACAAATTTCCCGAACCATTATATTTCTCTACATCAGAAAGGTATATAATCAT 1190  
Db |||||||  
4843 TTTTGGGAAACAAATTTCCCGAACCATTATATTTCTCTACATCAGAAAGGTATATAATCAT 4902  
Qy 1191 AAAAATCTTTTGAAGTCATTCTTTTACAGGAGTCCAAATACCAGAGAATGTTTATAGATACAC 1250  
Db |||||||  
4903 AAAAATCTTTTGAAGTCATTCTTTTACAGGAGTCCAAATACCAGAGAATGTTTATAGATACAC 4962  
Qy 1251 CATCAAAATTTGTAAGTGGCTTAATTTAGTTTATCACCCTTTGTCTCACTAAGAAAATAAATG 1310  
Db |||||||  
4963 CATCAAAATTTGTAAGTGGCTTAATTTAGTTTATCACCCTTTGTCTCACTAAGAAAATAAATG 5022  
Qy 1311 TGTACACAGTCTTAAAGCTGTATTTAGTTTATCACCCTTTGTCTCACTAAGAAAATAAATG 1370  
Db |||||||  
5023 TGTACACAGTCTTAAAGCTGTATTTAGTTTATCACCCTTTGTCTCACTAAGAAAATAAATG 5082  
Qy 1371 CAGGGTAAATTTATATCTCTCTTTTATGTTTGGGTATTAACACATATATCAATTT 1430  
Db |||||||  
5083 CAGGGTAAATTTATATCTCTCTTTTATGTTTGGGTATTAACACATATATCAATTT 5142  
Qy 1431 CTGTGTTTATCTAAAGTCGTTTCTGTTTCAAAATATGATTAATATCTCTTTTCTCT 1490  
Db |||||||  
5143 CTGTGTTTATCTAAAGTCGTTTCTGTTTCAAAATATGATTAATATCTCTTTTCTCT 5202  
Qy 1491 TCCAAATGTCTAAATCAATTTTATTAAGTTTCATTTGATATGCTTCCCTTAAATTTTATCT 1550  
Db |||||||  
5203 TCCAAATGTCTAAATCAATTTTATTAAGTTTCATTTGATATGCTTCCCTTAAATTTTATCT 5262  
Qy 1551 AAAGTCAATTTAGGAGGCTTACT 1610  
Db |||||||  
5263 AAAGTCAATTTAGGAGGCTTACT 5322  
Qy 1611 AGTCAATATTTACTGTAAACATAAATATATATTTTAAATAATATCCCACTTTATCAATTTTC 1670  
Db |||||||  
5323 AGTCAATATTTACTGTAAACATAAATATATATTTTAAATAATATCCCACTTTATCAATTTTC 5382  
Qy 1671 GTTTGTGTAACCTAAATGGGTCCTTTAGTTGAAGAAATAAGACCCACATTAATAAATGTCGTC 1730  
Db |||||||  
5383 GTTTGTGTAACCTAAATGGGTCCTTTAGTTGAAGAAATAAGACCCACATTAATAAATGTCGTC 5442  
Qy 1731 TTTTGTGTTTTTTTAAAGGATTTGAGCGTAGCGGAAAAATCCCTTTTCTTTCTTATCTTCAAT 1790  
Db |||||||  
5443 TTTTGTGTTTTTTTAAAGGATTTGAGCGTAGCGGAAAAATCCCTTTTCTTTCTTATCTTCAAT 5502  
Qy 1791 AATAAGGTAACCTATTGCGCGGATAGACTGTAAAC 1824  
Db |||||||  
5503 AATAAGGTAACCTATTGCGCGGAGGCTAGTTTAC 5536

RESULT 14  
US-10-032-393-53  
; Sequence 53, Application US/10032393  
; Publication No. US20030027286A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Wall, Daniel  
; APPLICANT: Gross, Molly  
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
; FILE REFERENCE: ELITRA.010A  
; CURRENT APPLICATION NUMBER: US/10/032.393  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/259,434  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: 09/948,993  
; PRIOR FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 53  
; LENGTH: 6852  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pXyl-T5-DD p15a  
US-10-032-393-53

Query Match 19.1%; Score 981.2; DB 15; Length 6852;  
Best Local Similarity 99.2%; Pred. No. 8.1e-181;  
Matches 986; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 831 TTAGAAAACCGACTGTAAGAGTACAGTCGGCATTATCTCATATTATATAAAGCCAGTCAT 890  
DB 4543 TTAGAAAACCGACTGTAAGAGTACAGTCGGCATTATCTCATATTATATAAAGCCAGTCAT 4602  
QY 891 TAGGCCTATCTGACAAATCTCGTAATGATAGAGTTCATTAACAATCCGTCATGATCAACATCAC 950  
DB 4603 TAGGCCTATCTGACAAATCTCGTAATGATAGAGTTCATTAACAATCCGTCATGATCAACATCAC 4662  
QY 951 AAACAGATGATGATACCTGTAAGATAGCGGTAAATATATTGAATACCTTTATTATGA 1010  
DB 4663 AAACAGATGATGATACCTGTAAGATAGCGGTAAATATATTGAATACCTTTATTATGA 4722  
QY 1011 ATTTTCTGCTGTAATAATGGGTAGAGGTAAATTAATTAATTAATTAATTAATTAATTAATTA 1070  
DB 4723 ATTTTCTGCTGTAATAATGGGTAGAGGTAAATTAATTAATTAATTAATTAATTAATTAATTA 4782  
QY 1071 ACCCAGTAATGAAGTCCGTAAGTAATAGAGGAAAGAAAGCAATTTTCAGGTATAGGTG 1130  
DB 4783 ACCCAGTAATGAAGTCCGTAAGTAATAGAGGAAAGAAAGCAATTTTCAGGTATAGGTG 4842  
QY 1131 TTTTGGGAAACCAATTTCCCGGAAACCAATTTATTTCTTACATCAGAAAGGTATAAATCAT 1190  
DB 4843 TTTTGGGAAACCAATTTCCCGGAAACCAATTTATTTCTTACATCAGAAAGGTATAAATCAT 4902  
QY 1191 AAACCTCTTTGAAGTCAATTTTACAGGAGTCCAAATACCAGAGATGTTTATAGATACAC 1250  
DB 4903 AAACCTCTTTGAAGTCAATTTTACAGGAGTCCAAATACCAGAGATGTTTATAGATACAC 4962  
QY 1251 CATCAAAATGATATAAGTGGCTCTAACTTATCCCAATAAAGCAATCTCCGTCGCTAT 1310  
DB 4963 CATCAAAATGATATAAGTGGCTCTAACTTATCCCAATAAAGCAATCTCCGTCGCTAT 5022  
QY 1311 TGTAACCAAGTCTTAAAGCTGTATTTAGTTATACACCTTGCTACTAAGAAATAAATG 1370  
DB 5023 TGTAACCAAGTCTTAAAGCTGTATTTAGTTATACACCTTGCTACTAAGAAATAAATG 5082  
QY 1371 CAGGGTAAATTTATATCTCTTGTGTTTATGTTTCGGTATAAAGCACTAATATCAATTT 1430  
DB 5083 CAGGGTAAATTTATATCTCTTGTGTTTATGTTTCGGTATAAAGCACTAATATCAATTT 5142

RESULT 15  
US-10-466-642-24/c  
; Sequence 24, Application US/10466642  
; Publication No. US20040048343A1  
; GENERAL INFORMATION:  
; APPLICANT: OmniGene BioProducts  
; TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-  
; TITLE OF INVENTION: HYDROXY-3-METHYL-BUTYRLAMINO)-PROPIONIC ACID (HMBPA)  
; FILE REFERENCE: BGI-1465C  
; CURRENT APPLICATION NUMBER: US/10/466.642  
; CURRENT FILING DATE: 2003-07-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 3867  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: vector  
US-10-466-642-24

Query Match 19.0%; Score 978.4; DB 13; Length 3867;  
Best Local Similarity 99.9%; Pred. No. 2.1e-181;  
Matches 979; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 831 TTAGAAAACCGACTGTAAGAGTACAGTCGGCATTATCTCATATTATATAAAGCCAGTCAT 890  
DB 2460 TTAGAAAACCGACTGTAAGAGTACAGTCGGCATTATCTCATATTATATAAAGCCAGTCAT 2401  
QY 891 TAGGCCTATCTGACAAATCTCGTAATGATAGAGTTCATTAACAATCCGTCATGATCAACATCAC 950  
DB 2400 TAGGCCTATCTGACAAATCTCGTAATGATAGAGTTCATTAACAATCCGTCATGATCAACATCAC 2341  
QY 951 AAACAGATGATGATACCTGTAAGATAGCGGTAAATATATTGAATACCTTTATTATGA 1010  
DB 2340 AAACAGATGATGATACCTGTAAGATAGCGGTAAATATATTGAATACCTTTATTATGA 2281  
QY 1011 ATTTTCTGCTGTAATAATGGGTAGAGGTAAATTAATTAATTAATTAATTAATTAATTAATTA 1070  
DB 2280 ATTTTCTGCTGTAATAATGGGTAGAGGTAAATTAATTAATTAATTAATTAATTAATTAATTA 2221  
QY 1071 ACCCAGTAAATGAAGTCCATGGAATATAGAAAGAGAAAGCAATTTTCAGGTATAGGTG 1130  
DB 2220 ACCCAGTAAATGAAGTCCATGGAATATAGAAAGAGAAAGCAATTTTCAGGTATAGGTG 2161



QY 1131 TTTTGGGAAACAAATTTCCCGAACCAATATATTTCTCTACATCAGAAAGGTATAAATCAT 1190  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2160 TTTTGGGAAACAAATTTCCCGAACCAATATATTTCTCTACATCAGAAAGGTATAAATCAT 2101  
QY 1191 AAAACCTCTTGAAGTCATCTTTACAGGAGTCCAAATACACAGAGAAATGTTTTGATACAC 1250  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2100 AAAACCTCTTGAAGTCATCTTTACAGGAGTCCAAATACACAGAGAAATGTTTTGATACAC 2041  
QY 1251 CATCAAAAATTTGATAAAGTGGCTCTAACTTTATCCCAATACCTAACCTCCGTCGCTAT 1310  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2040 CATCAAAAATTTGATAAAGTGGCTCTAACTTTATCCCAATACCTAACCTCCGTCGCTAT 1981  
QY 1311 TGTACACAGTTCTAAAGCTGATTTGAGTTTATCACCCCTGTCTACAGAAATTAATG 1370  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1980 TGTACACAGTTCTAAAGCTGATTTGAGTTTATCACCCCTGTCTACAGAAATTAATG 1921  
QY 1371 CAGGTAATAATTTATATCTTTCTTTTATGTTTCGTTATATAAAACACTAATATCAATTT 1430  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1920 CAGGTAATAATTTATATCTTTCTTTTATGTTTCGTTATATAAAACACTAATATCAATTT 1861  
QY 1431 CTGTGGTTATATAAAGTCGTTTGTGGTTCAATAATGATTAATAATCTCTTTCTCT 1490  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1860 CTGTGGTTATATAAAGTCGTTTGTGGTTCAATAATGATTAATAATCTCTTTCTCT 1801  
QY 1491 TCCAATTTGCTAAATCAATTTTATTAAGTTTCAATTTGATATGCTCTTAAATTTTATCT 1550  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1800 TCCAATTTGCTAAATCAATTTTATTAAGTTTCAATTTGATATGCTCTTAAATTTTATCT 1741  
QY 1551 AAAGTGAATTTAGGAGGCTTACTTGTCTGCTTTCTTCAATAGAAATCAATCTTTTAA 1610  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1740 AAAGTGAATTTAGGAGGCTTACTTGTCTGCTTTCTTCAATAGAAATCAATCTTTTAA 1681  
QY 1611 AGTCAATATTACTGTAACATAATATATTTTAAATATATCCACCTTTATCCAAATTTTC 1670  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1680 AGTCAATATTACTGTAACATAATATATTTTAAATATATCCACCTTTATCCAAATTTTC 1621  
QY 1671 GTTCTTGAACCTAATGGTGTCTTTAGTTGAAGAAATAAGACCAATTAATAAATGTGGTC 1730  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1620 GTTCTTGAACCTAATGGTGTCTTTAGTTGAAGAAATAAGACCAATTAATAAATGTGGTC 1561  
QY 1731 TTTTGTGTTTTTTTAAAGGATTTGAGGCTAGCGAAATCCCTTTCTTTCTTATCTTGAT 1790  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1560 TTTTGTGTTTTTTTAAAGGATTTGAGGCTAGCGAAATCCCTTTCTTTCTTATCTTGAT 1501  
QY 1791 AATAAGGTAACCTATTGGCG 1810  
DB |||||||||||||||||||  
1500 AATAAGGTAACCTATTGGCG 1481

Search completed: September 24, 2004, 19:24:45  
Job time : 1470.26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 23, 2004, 17:04:19 ; Search time 7801.56 Seconds  
(without alignments).  
19682.126 Million cell updates/sec

Title: US-10-030-390-1  
Perfect score: 5142  
Sequence: 1 gaattcagctcggtaccg.....ccgtctgtagcttccttaag 5142

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estrov:\*

6: em\_estrpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_nam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_fod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	200.2	3.9	456	9 AA060105	AA060105 mj71f11.r
2	200.2	3.9	456	14 WB3072	WB3072 mf09c04.r1
3	198.6	3.9	354	13 BX102466	BX102466 BX102466
4	198.6	3.9	390	9 AI323443	AI323443 mj71f11.x

5	198.6	3.9	451	9 AI893383	AI893383 mk16b08.y
6	198.6	3.9	455	9 AI323089	AI323089 mj71f11.y
7	198.6	3.9	473	13 BY708773	BY708773 BY708773
8	198.6	3.9	474	11 AK008625	AK008625 Mus muscu
9	198.6	3.9	474	11 AK010179	AK010179 Mus muscu
10	198.6	3.9	475	11 AK008809	AK008809 Mus muscu
11	198.6	3.9	477	11 AK008873	AK008873 Mus muscu
12	198.6	3.9	486	13 BY708705	BY708705 BY708705
13	198.6	3.9	505	13 BY708809	BY708809 BY708809
14	198.6	3.9	994	9 AV072530	AV072530 AV072530
15	195.4	3.8	384	13 BY076652	BY076652 BY076652
16	195.4	3.8	354	13 BX101861	BX101861 BY101861
17	185.8	3.6	311	10 BB564851	BB564851 BB564851
18	182.6	3.6	435	9 AA097254	AA097254 mk16b08.r
19	181	3.5	311	10 BB564891	BB564891 BB564891
20	180.2	3.5	443	14 W12684	W12684 ma52b04.r1
21	165	3.2	238	29 AV419513	AV419513 Mus muscu
22	163	3.2	307	10 BB565090	BB565090 BB565090
23	161.2	3.1	306	10 BB565053	BB565053 BB565053
24	156.4	3.0	332	9 AV076259	AV076259 AV076259
C 25	155	3.0	705	10 BB667664	BB667664 BB667664
C 26	155	3.0	1515	11 AK050258	AK050258 Mus muscu
27	154.8	3.0	263	10 BB564847	BB564847 BB564847
28	148.4	2.9	300	10 BB564905	BB564905 BB564905
29	141	2.7	421	10 BF738530	BF738530 BF738530
C 30	137.8	2.7	354	10 BF707544	BF707544 AI43.LE.A
31	131.8	2.6	264	10 BB564882	BB564882 BB564882
32	130.8	2.5	298	10 BB565023	BB565023 BB565023
33	130.8	2.5	812	12 BM536445	BM536445 5'AD3.Vad
34	130.4	2.5	602	28 B06973	B06973 CSRL-8b6-t
35	127.4	2.5	242	10 BB565149	BB565149 BB565149
36	125.2	2.4	255	10 BB564787	BB564787 BB564787
37	124.8	2.4	317	9 AV075620	AV075620 AV075620
38	122.6	2.4	320	9 AV081837	AV081837 AV081837
39	120.4	2.3	254	10 BB565113	BB565113 BB565113
40	120.2	2.3	499	29 TA301803Q	TA301803Q T. brucei
41	120.2	2.3	550	29 TA144805Q	TA144805Q T. brucei
42	119.8	2.3	304	9 AV075572	AV075572 AV075572
43	118	2.3	503	29 TA344H03Q	TA344H03Q T. brucei
44	117.4	2.3	653	29 TA371H07Q	TA371H07Q T. brucei
45	117.4	2.3	659	29 TA371A11Q	TA371A11Q T. brucei

ALIGNMENTS

RESULT 1  
AA060105  
LOCUS  
DEFINITION  
mj71f11.r1 Soates mouse p3NNF19.5 Mus musculus cDNA clone  
IMAGE:481591 5' similar to gb:221858 M.Musculus mRNA for p domain  
protein (MOUSE);, mRNA sequence.  
AA060105  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 456)  
Marra.M., Hillier.L., Allen.M., Bowles.M., Dietrich.N., Dubuque.T.,  
Gaisel.S., Kucaba.T., Lacy.M., Le.M., Martin.J., Morris.M.,  
Schellenberg.K., Steptoe.M., Tan.F., Underwood.K., Moore.B.,  
Theising.B., Wylie.T., Lennon.G., Soares.B., Wilson.R. and  
Waterston,R.  
The WashU-HMNI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:292325  
Putative full length read  
vector to vector length is 457  
Seq primer: -28M13 rev2 from Amersham.  
Location/Qualifiers  
1. .456

## FEATURES

source  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:481581"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares mouse p3NM19.5"  
/note="Vector: p7T3D (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5],  
TGTACCAATCTGAAGTGGAGCGCGCATTTTCTTTT 3',  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified p7T3 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M.Patima Bonaldo. RNA was kindly provided by  
Dr. Minoru Ko (Wayne State University)."

## ORIGIN

Query Match 3.9%; Score 200.2; DB 9; Length 456;  
Best Local Similarity 88.6%; Pred. No. 5.4e-26;  
Matches 217; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 153 TGTCTACAGTCATATTCTGCTGCAGCCCGTGTTCAGGTGTTTACGCCAGGCCCCAGG 212  
DB |||||  
31 TGTGTCCTCGCTGTGGTCTCATGTGGCCTTCGAGCCTTCGCCAGGCCCCAGG 90  
QY 213 CCCAGGCCAGGAAGAACATGTATCATGGCCCCCGGGAGAGGATAAATGTGCTTCC 272  
DB |||||  
91 CCCAGGCCAGGAAGAACATGTATCATGGCCCCCGGGAGAGGATAAATGTGCTTCC 150  
QY 273 CCGGTGTACCGCCAGCAGTGCACGGAGAGGTGCTGTTTTCATCAGAGTCTCCGG 332  
DB |||||  
151 CCGGTGTACCGCCAGCAGTGCACGGAGAGGTGCTGTTTTCATCAGAGTCTCCGG 210  
QY 333 GATTCCCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGAGAGAAATGTCCT 392  
DB |||||  
211 GATTCCCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGAGAGAAATGTCCT 270

QY 393 TCTAA 397  
DB |||||  
271 TCTAA 275

RESULT 2  
W83072 456 bp mRNA linear EST 12-SEP-1996  
LOCUS m09c04.r1 Soares mouse p3NM19.5 Mus musculus cDNA clone  
DEFINITION IMAGE:404550 5', similar to gb:221858 M.Musculus mRNA for P domain  
protein (MOUSE);, mRNA sequence.

W83072  
W83072.1 GI:1540299

ACCESSION  
VERSION  
KEYWORDS  
SOURCE

Mus musculus (house mouse)  
EST.

ORGANISM  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS

1 (bases 1 to 456)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Willson, R. and  
Waterston, R.

TITLE  
JOURNAL  
COMMENT

The WashU-HMI Mouse EST Project  
Unpublished (1996)  
On Sep 12, 1996 this sequence version replaced gi:1394024.  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:248318  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 321.  
Location/Qualifiers  
1. .456

FEATURES  
source

/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:404550"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares mouse p3NM19.5"  
/note="Vector: p7T3D (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5],  
TGTACCAATCTGAAGTGGAGCGCGCATTTTCTTTT 3',  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified p7T3 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M.Patima Bonaldo. RNA was kindly provided by  
Dr. Minoru Ko (Wayne State University)."

## ORIGIN

Query Match 3.9%; Score 200.2; DB 14; Length 456;  
Best Local Similarity 88.6%; Pred. No. 5.4e-26;  
Matches 217; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 153 TGTCTACAGTCATATTCTGCTGCAGCCCGTGTTCAGGTGTTTACGCCAGGCCCCAGG 212  
DB |||||  
31 TGTGTCCTCGCTGTGGTCTCATGTGGCCTTCGAGCCTTCGCCAGGCCCCAGG 90  
QY 213 CCCAGGCCAGGAAGAACATGTATCATGGCCCCCGGGAGAGGATAAATGTGCTTCC 272  
DB |||||  
91 CCCAGGCCAGGAAGAACATGTATCATGGCCCCCGGGAGAGGATAAATGTGCTTCC 150  
QY 273 CCGGTGTACCGCCAGCAGTGCACGGAGAGGTGCTGTTTTCATCAGAGTCTCCGG 332  
DB |||||  
151 CCGGTGTACCGCCAGCAGTGCACGGAGAGGTGCTGTTTTCATCAGAGTCTCCGG 210  
QY 333 GATTCCCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGAGAGAAATGTCCT 392  
DB |||||  
211 GATTCCCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGAGAGAAATGTCCT 270

RESULT 3  
BY102466  
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BY102466 354 bp mRNA linear EST 07-DEC-2002  
BY102466 RIKEN full-length tissues, pooled tissues, adult spleen,  
etc. Mus musculus cDNA clone K630145D13 5', mRNA sequence.  
BY102466  
BY102466.1 GI:26213083  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



This clone was previously sequenced on the 5' end only, this new data is from the 3' end  
High quality sequence stop: 376.

# FEATURES

Location/Qualifiers  
1..390  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:481581"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares mouse p3NMf19.5"  
/note="Vector: p7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5, TGTACCAATCTGAAGTGGAGCGCCGCAATTTTITTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

# ORIGIN

Query Match 3.9%; Score 198.6; DB 9; Length 390;  
Best Local Similarity 98.0%; Pred. No. 1.1e-25;  
Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 193 TGTTTACGCCAGGCCAGGCCAGGAGAAACATGTATCATGCCCCCGGGA 252  
DB 384 TGCCAGAGCCAGGCCAGGCCAGGAGAAACATGTATCATGCCCCCGGGA 325  
QY 253 GAGGATAAATTGTGCTTCCCGGTGTACCGCCAGCAGTGACGAGAGGTTCTG 312  
DB 324 GAGGATAAATTGTGCTTCCCGGTGTACCGCCAGCAGTGACGAGAGGTTCTG 265  
QY 313 TTTTATCAGAGTGTCGGGATTCCTTCTAA 397  
DB 264 TTTTATCAGAGTGTCGGGATTCCTTCTAA 180  
QY 373 TCAGAAGAAGATGTCCTTCTAA 397  
DB 204 TCAGAAGAAGATGTCCTTCTAA 180

RESULT 5  
AI893383  
LOCUS  
DEFINITION  
mk16b08.y1 Soares mouse p3NMf19.5 Mus musculus cDNA clone  
IMAGE:493047 5', similar to gb:221858 M.Musculus mRNA for P domain  
protein (MOUSE);, mRNA sequence.

ACCESSION  
AI893383  
VERSION  
AI893383.1 GI:5599285  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 451)  
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,  
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,  
Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,  
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,  
Waterston,R. and Wilson,R.  
The WashU-NCI Mouse EST Project 1999  
Unpublished (1999)  
Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu

TITLE  
JOURNAL  
COMMENT  
The WashU-HHMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
This read is a RESEQUENCE of a previously sequenced mouse clone  
This read has been verified (found to hit its original self in the correct orientation)

MGI:296495  
Seq primer: -40RP from Gibco  
High quality sequence stop: 448.  
Location/Qualifiers  
1..451  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:493047"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares mouse p3NMf19.5"  
/note="Vector: p7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5, TGTACCAATCTGAAGTGGAGCGCCGCAATTTTITTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

# FEATURES

Location/Qualifiers  
1..451  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:493047"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares mouse p3NMf19.5"  
/note="Vector: p7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5, TGTACCAATCTGAAGTGGAGCGCCGCAATTTTITTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

# ORIGIN

Query Match 3.9%; Score 198.6; DB 9; Length 451;  
Best Local Similarity 98.0%; Pred. No. 1.1e-25;  
Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 193 TGTTTACGCCAGGCCAGGCCAGGAGAAACATGTATCATGCCCCCGGGA 252  
DB 66 TGCCAGAGCCAGGCCAGGCCAGGAGAAACATGTATCATGCCCCCGGGA 125  
QY 253 GAGGATAAATTGTGCTTCCCGGTGTACCGCCAGCAGTGACGAGAGGTTCTG 312  
DB 126 GAGGATAAATTGTGCTTCCCGGTGTACCGCCAGCAGTGACGAGAGGTTCTG 185  
QY 313 TTTTATCAGAGTGTCGGGATTCCTTCTAA 451  
DB 186 TTTTATCAGAGTGTCGGGATTCCTTCTAA 270  
QY 373 TCAGAAGAAGATGTCCTTCTAA 451  
DB 246 TCAGAAGAAGATGTCCTTCTAA 270

# RESULT 6

AI323089  
LOCUS  
DEFINITION  
mj7f11.y1 Soares mouse p3NMf19.5 Mus musculus cDNA clone  
IMAGE:481581 5', similar to gb:221858 M.Musculus mRNA for P domain  
protein (MOUSE);, mRNA sequence.

ACCESSION  
AI323089  
VERSION  
AI323089.1 GI:4057518  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 455)  
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
The WashU-HHMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project

# TITLE JOURNAL COMMENT

The WashU-HHMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project

WASHU-HHMI Mouse EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:292325

This read is a RESEQUENCE of a previously sequenced mouse clone  
correct orientation)  
putative full length read  
vector to vector length is 457  
Seq primer: -40RP from Gibco  
High quality sequence stop: 439.  
Location/Qualifiers

## FEATURES

source

```

1. 455
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:481581"
/dev_host="DHI0B (ampicillin resistant)"
/lab_host="DHI0B (ampicillin resistant)"
/clone_lib="Seares mouse p3MMP19.5"
/notes="Vector: p773D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Scares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."
```

## ORIGIN

```

Query Match      3.9%; Score 198.6; DB 9; Length 455;
Best Local Similarity 98.0%; Pred. No. 1.1e-25;
Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 193 TGTTTACGCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCATGTCGCCCGGGA 252
DB 72 TGCCCCAGGCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCATGTCGCCCGGGA 131
QY 253 GAGGATTAATTTGCTTCCCGGTTCCACCCAGCAGTCACGAGAGAGGTTGCTG 312
DB 132 GAGGATTAATTTGCTTCCCGGTTCCACCCAGCAGTCACGAGAGAGGTTGCTG 191
QY 313 TTTTGTATGACAGTGTCCGGGATTCCTGTTGCTTCCACCCATGCGCAGAACAC 372
DB 192 TTTTGTATGACAGTGTCCGGGATTCCTGTTGCTTCCACCCATGCGCAGAACAC 251
QY 373 TCAAGAAGAAGATGTCCTTCTAA 397
DB 252 TCAAGAAGAAGATGTCCTTCTAA 276
```

## RESULT 7

```

BY708773
LOCUS BY708773 RIKEN full-length enriched, adult male stomach Mus
DEFINITION musculus cDNA clone 2210409105 5', mRNA sequence.
ACCESSION BY708773
VERSION BY708773.1 GI:271119967
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 473)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
```

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Moganai, A., Schonbach, C., Gojibori, T., Baldarelli, R., Hill, D.P., Buit, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gusincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perlea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shinada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, E., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

## TITLE

Analysis of the mouse transcriptome based on functional annotation

## JOURNAL

NATURE 420, 563-573 (2002)

## MEDLINE

22354683

## PUBMED

12466851

## COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Sueniro-cho, Isurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp  
URL: http://genome.gsc.riken.go.jp/  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN,  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

## FEATURES

source

```

1. 473
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="2210409105"
/sex="male"
```

```

/tissue type="stomach"
/dev stage="adult"
/lab host="SORL"
/clone_lib="RIKEN full-length enriched, adult male
stomach"
/note="Site 1: XhoI; Site 2: SstI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'.
GAGAGAGAGCGCCGCACTCGAGTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'.
GAGAGAGATTCGCGATTATTAATTAATCCCCCCCCCC 3']. cDNA
was cleaved with XhoI and SstI."

```

## ORIGIN

```

Query Match          3.9%; Score 198.6; DB 13; Length 473;
Best Local Similarity 98.0%; Pred. No. 1e-25;
Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 193 TGTTCACCCAGGCCAGGCCAGGCCAGGAAACATGTATCATATGCCCCCGGGA 252
    |||
DB 89 TGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGCCCCCGGGA 148

QY 253 GAGGATAAATGTGCTTCCCGGTGTCCAGCCCGCAGCGTGCAGGAGAGGTTCTG 312
    |||
DB 149 GAGGATAAATGTGCTTCCCGGTGTCCAGCCCGCAGCGTGCAGGAGAGGTTCTG 208

QY 313 TTTGATCAGATGTCGCGGATTCCTCGTGTCTTCCACCCCATGCCATCGAGAAC 372
    |||
DB 209 TTTGATCAGATGTCGCGGATTCCTCGTGTCTTCCACCCCATGCCATCGAGAAC 268

QY 373 TCAGAAGAAGAAATGCTCTCTAA 397
    |||
DB 269 TCAGAAGAAGAAATGCTCTCTAA 293

```

## RESULT 8

```

AK008625          474 bp mRNA linear HTC 20-SEP-2003
LOCUS             Mus musculus adult male stomach cDNA, RIKEN full-length enriched
DEFINITION        library, clone:2200002M16 product:trefol factor 1, full insert
sequence.
ACCESSION         AK008625
VERSION           AK008625.1 GI:12842925
KEYWORDS          HTC; CAP trapper.
SOURCE            Mus musculus (house mouse)
ORGANISM          Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

## REFERENCE

```

1. Carninci, P. and Hayashizaki, Y.
   High-efficiency full-length cDNA cloning
   Mech. Enzymol. 303, 19-44 (1999)
99279253          PUBMED
10349636          PUBMED
2. Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
   Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
   Normalization and subtraction of cap-trapper-selected cDNAs to
   prepare full-length cDNA libraries for rapid discovery of new genes
   Genome Res. 10 (10), 1617-1630 (2000)
20499374          PUBMED
11042159          PUBMED
3. Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
   Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
   Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

```

```

TITLE            The RIKEN Genome Exploration Research Group Phase II Team and the
JOURNAL           FANTOM Consortium.
MEDLINE           Functional annotation of a full-length mouse cDNA collection
PUBMED           Nature 409, 685-690 (2001)
REFERENCE         5. The FANTOM Consortium and the RIKEN Genome Exploration Research
                  Group Phase I & II Team.
                  Analysis of the mouse transcriptome based on functional annotation
                  of 60,770 full-length cDNAs
                  Nature 420, 563-573 (2002)
JOURNAL           Nature 420, 563-573 (2002)
AUTHORS           6 (bases 1 to 474)
                  Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
                  Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
                  Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
                  Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
                  Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
                  Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
                  Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
                  Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
                  Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
                  Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
                  Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
                  Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
                  Direct Submission
                  Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
                  Physical and Chemical Research (RIKEN), Laboratory for Genome
                  Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
                  RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
                  Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,
                  URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
                  Fax: 81-45-503-9216]
                  Please visit our web site (http://genome.gsc.riken.go.jp/) for
                  further details.
                  cDNA library was prepared and sequenced in Mouse Genome
                  Encyclopedia Project of Genome Exploration Research Group in Riken
                  Genomic Sciences Center and Genome Science Laboratory in RIKEN.
                  Division of Experimental Animal Research in Riken contributed to
                  prepare mouse tissues. First strand cDNA was primed with a primer
                  [5'. GAGAGAGAGAGGATCCAGAGCTTTTTTTTTTTTTTNN 3'], cDNA was
                  prepared by using trehalose thermo-activated reverse transcriptase
                  and subsequently enriched for full-length by cap-trapper. Second
                  strand cDNA was prepared with the primer adapter of sequence [5'.
                  GAGAGAGATTCGCGATTATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved
                  with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
                  Host: SOLR.
FEATURES          1. Location/Qualifiers
                  location=1..474
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="C57BL/6J"
                  /db_xref="FANTOM DB:2200002M16"
                  /db_xref="MGI:1898509"
                  /db_xref="taxon:10090"
                  /clone="2200002M16"
                  /sex="male"
                  /tissue type="stomach"
                  /clone_lib="RIKEN full-length enriched mouse cDNA library"
                  /dev stage="adult"
                  31..294
                  /note="unnamed protein product; putative
                  trefol factor 1 (MGI:189135)"
                  /codon_start=1
                  /protein_id="BAB25787.1"

```

## CDS

```

1..474
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:2200002M16"
/db_xref="MGI:1898509"
/db_xref="taxon:10090"
/clone="2200002M16"
/sex="male"
/tissue type="stomach"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="adult"
31..294
/note="unnamed protein product; putative
trefol factor 1 (MGI:189135)"
/codon_start=1
/protein_id="BAB25787.1"

```

```

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipipillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

```

## AUTHORS

```

4. The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
20530913
11076861

```

## AUTHORS

```

5. The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

```

## AUTHORS

```

6 (bases 1 to 474)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'. GAGAGAGAGGATCCAGAGCTTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5'.
GAGAGAGATTCGCGATTATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

```

## AUTHORS

```

TITLE            The RIKEN Genome Exploration Research Group Phase II Team and the
JOURNAL           FANTOM Consortium.
MEDLINE           Functional annotation of a full-length mouse cDNA collection
PUBMED           Nature 409, 685-690 (2001)
REFERENCE         5. The FANTOM Consortium and the RIKEN Genome Exploration Research
                  Group Phase I & II Team.
                  Analysis of the mouse transcriptome based on functional annotation
                  of 60,770 full-length cDNAs
                  Nature 420, 563-573 (2002)
JOURNAL           Nature 420, 563-573 (2002)
AUTHORS           6 (bases 1 to 474)
                  Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
                  Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
                  Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
                  Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
                  Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
                  Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
                  Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
                  Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
                  Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
                  Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
                  Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
                  Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
                  Direct Submission
                  Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
                  Physical and Chemical Research (RIKEN), Laboratory for Genome
                  Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
                  RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
                  Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,
                  URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
                  Fax: 81-45-503-9216]
                  Please visit our web site (http://genome.gsc.riken.go.jp/) for
                  further details.
                  cDNA library was prepared and sequenced in Mouse Genome
                  Encyclopedia Project of Genome Exploration Research Group in Riken
                  Genomic Sciences Center and Genome Science Laboratory in RIKEN.
                  Division of Experimental Animal Research in Riken contributed to
                  prepare mouse tissues. First strand cDNA was primed with a primer
                  [5'. GAGAGAGAGGATCCAGAGCTTTTTTTTTTTTTTNN 3'], cDNA was
                  prepared by using trehalose thermo-activated reverse transcriptase
                  and subsequently enriched for full-length by cap-trapper. Second
                  strand cDNA was prepared with the primer adapter of sequence [5'.
                  GAGAGAGATTCGCGATTATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved
                  with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
                  Host: SOLR.

```

## COMMENT

```

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'. GAGAGAGAGGATCCAGAGCTTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5'.
GAGAGAGATTCGCGATTATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

```

## FEATURES

```

1. Location/Qualifiers
location=1..474
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:2200002M16"
/db_xref="MGI:1898509"
/db_xref="taxon:10090"
/clone="2200002M16"
/sex="male"
/tissue type="stomach"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="adult"
31..294
/note="unnamed protein product; putative
trefol factor 1 (MGI:189135)"
/codon_start=1
/protein_id="BAB25787.1"

```



```

/db xref="GI:112842926"
/translations="MEKVTVCVLAVLMLAFGLAQQAQAQAQAEETCINWPRERINC
GFPQVTAQQCTERGCGCFDDSVRGFPFPHFPAIENTOEBCPF"

ORIGIN

Query Match          3.9%  Score 198.6;  DB 11;  Length 474;
Best Local Similarity 98.0%;  Pred. No. 1e-25;
Matches 201;  Conservative 0;  Mismatches 4;  Indels 0;  Gaps 0;

QY 193 TGTTTACGCCAGGCCCCAGGCCAGGCCAGGAGAAACAATGTATCATGCCCCCGGGA 252
Db 90 TGCCAGGCCAGGCCCCAGGCCAGGCCAGGAGAAACAATGTATCATGCCCCCGGGA 149

QY 253 GAGGATAAATGTGGCTTCCCGGTGTACCGCCGACAGTGCACGAGAGAGGTTGCTG 312
Db 150 GAGGATAAATGTGGCTTCCCGGTGTACCGCCGACAGTGCACGAGAGAGGTTGCTG 209

QY 313 TTTTGTATGACAGTGTCCGGGATTCCTGGTGTCTTCCACCCCATGCGCAGAACAC 372
Db 210 TTTTGTATGACAGTGTCCGGGATTCCTGGTGTCTTCCACCCCATGCGCAGAACAC 269

QY 373 TCAGAAGAAGAAATGTCCTTCTAA 397
Db 270 TCAGAAGAAGAAATGTCCTTCTAA 294

RESULT 9
AKO10179
LOCUS
DEFINITION
Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310075107 product:trefoil factor 1, full insert
sequence.
ACCESSION
AKO10179
VERSION
AKO10179.1 GI:12845437
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, K., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watahiki, M.,
Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Yoneda, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research

```

Yoshino, M., Muranatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
Please visit our web site (http://genome.gsc.riken.go.jp/) for  
further details.

CDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGAGAGAAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. Second  
strand cDNA was prepared with the primer adapter of sequence [5'  
GAGAGAGATTTCTCGAGTAAATTAATTAATTCCTCCCTCCCTCC 3']. cDNA was cleaved  
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

Location/Qualifiers

1..475	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="CS7BL/6J"
	/db_xref="FANTOM_DB:2210403C21"
	/db_xref="MGI:1895586"
	/db_xref="taxon:10090"
	/clone="2210403C21"
	/sex="male"
	/tissue_type="stomach"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
31..294	/dev_stage="adult"

CDS

1..294	/note="unnamed protein product; putative
	trefoil factor 1 (MGD MGI:88135)"
	/codon_start=1
	/protein_id="BAB95907.1"
	/db_xref="GI:128432293"
	/translation="MEHKVICVLAVLMLAFGSLAQAAQAAQAEETCIMAPRERINC
	GFPGVTAQQCTERCCCFDDSVRGFPFWCFHPMAIENTQEEQFF"
457..462	/note="putative"

polyA\_signal

polyA\_site

475	/note="putative"
-----	------------------

ORIGIN

Query Match	3.9%;	Score 198.6;	DB 11;	Length 475;
Best Local Similarity	98.0%;	Pred. No. le-25;		
Matches 201;	Conservative	0;	Mismatches 4;	Indels 0;
Gaps	0;			

QY 193 TGTATTACCCAGGCCAGGCCAGGCCAGGAAGAAACATGATCATGGCCCCCGGGA 252

DB 90 TGCCAGGCCAGGCCAGGCCAGGCCAGGAAGAAACATGATCATGGCCCCCGGGA 149

QY 253 GAGGATAAATTGTGGCTTCCCGGTGTTCACCGCCAGCAGTGCACGGAGAGAGTGTCTG 312

DB 150 GAGGATAAATTGTGGCTTCCCGGTGTTCACCGCCAGCAGTGCACGGAGAGAGTGTCTG 209

QY 313 TTTTCATCAGAGTGTCCGGGATTCCTGGTGTCTTCACCCCATGGCCATCGAAGACAC 372

DB 210 TTTTCATCAGAGTGTCCGGGATTCCTGGTGTCTTCACCCCATGGCCATCGAAGACAC 269

QY 373 TCAAGAGAAGAAATGTCCCTTCTAA 397

DB 270 TCAAGAGAAGAAATGTCCCTTCTAA 294

RESULT 11  
AK008873



Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schönbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Baralov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusic, V., Chochoia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Glissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierksi, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Portea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlested, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kisukawa, T., Konno, H., Nakamura, M., Sakazume, N., Satoh, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y., Itch, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasuniishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466951

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-2 Suenhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsic.riken.go.jp,  
 URU: http://genome.gsc.riken.go.jp/  
 Adachi, J., Akawa, K., Akimura, T., Arakawa, T., Carninci, P.,  
 Fukutani, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,  
 Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,  
 Komatsu, S., Konno, H., Koya, S., Miyazaki, A., Murata, R., Nakamura, M.,  
 Nondura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,  
 Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,  
 Takeda, Y., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: DNA time sequence Catalogue [2] 281-289 (2001). A cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

```

FEATURES
source
1. 486
/organism="Mus muscu
/mol_type="mRNA"
/db_xref="taxon:100
/clonseq="2210403C21"
sex="male"

```

```

/tissue_type="stomach"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male
stomach"
/note="site_1: XhoI; Site_2: SstI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGACGGCGCGCACTCGAGTTTTTTTTTTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCC 3']. cDNA
was cleaved with XhoI and SstI."

```

## ORIGIN

Query Match	3.9%;	Score 198.6;	DB 13;	Length 486;
Best Local Similarity	98.0%;	Pred. No. 1e-25;		
Matches 201;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	193	TGTTTACGCCCCAGGCCCCAGGCCCCAGGAGAAACATGTATCATGGCCCCCGGGA	252	
Db	90	TGCCACAGGCCCCAGGCCCCAGGCCCCAGGAGAAACATGTATCATGGCCCCCGGGA	149	
QY	253	GAGGATAAATTGTGGCTTCCCGGTGTCCACGCCCAGCAGTGCACGGAGAGAGTTTCGTG	312	
Db	150	GAGGATAAATTGTGGCTTCCCGGTGTCCACGCCCAGCAGTGCACGGAGAGAGTTTCGTG	209	
QY	313	TTTTTGATGACAGTGTCCGGGGAAATCCCGTGTGCTTTCCACCCCATGGCCCATCGAAGACAC	372	
Db	210	TTTTTGATGACAGTGTCCGGGGAAATCCCGTGTGCTTTCCACCCCATGGCCCATCGAGACAC	269	
QY	373	TCAGAAGAAGAAATGTCCCTTCTAA	397	
Db	270	TCAGAAGAAGAAATGTCCCTTCTAA	294	

RESULT 13

BY708809	LOCUS	BY708809	505 bp	mRNA	linear	EST 16-DEC-2002
BY708809	DEFINITION	RIKEN full-length enriched, adult male stomach Mus musculus cDNA clone 2210414A03 5', mRNA sequence.				
BY708809	ACCESSION	BY708809				
BY708809	VERSION	1	GI:27120003			
BY708809	KEYWORDS	EST.				

SOURCE	ORGANISM
Mus musculus (house mouse)	Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE  
AUTHORS

1. (bases 1 to 505)  
Otake, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
Kiyosawa, H., Taga, K., Tomaru, Y., Hasegawa, Y., Wogami, A.,  
Shionobuchi, C., Gojobori, T., Baldarelli, R., Hill, D. P., Sult, C.,  
Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,  
Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusaica, V.,  
Chothia, C., Corbani, L. E., Cousins, S., Dalla, S., Dragani, I. A.,  
Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,  
Gariboldi, M., Giesli, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustinich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kenzierski, R. M., King, B. L., Konagaya, A.,  
Kurochkin, I. V., Lee, Y., Ledzierski, B., Lyons, P. A., Maglott, D. R.,  
Mallais, L., Marchionni, L., Mckenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,  
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Simple, C. A., Setou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,





adult,tissue\_type=amnion,sex=female),(dev\_stage=10 days  
neonate,tissue\_type=brain,sex=mix),(dev\_stage=10 days  
neonate,tissue\_type=thymus,sex=mix),(dev\_stage=10 days  
neonate,tissue\_type=heart,sex=mix)"

ORIGIN

Query Match	3.8%;	Score 195.4;	DB 13;	Length 384;
Best Local Similarity	97.1%;	Fred. No. 4.2e-25;		
Matches 199;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	193	TCGTTTACGCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCATGGCCCCCGGGA	252	
Db	86	TGCCCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCATGGCCCCCGGGA	145	
QY	253	GAGGATAAATTGTGGCTTCCCGGTGTACCGCCAGCAGTGCACGGAGAGAGTTGCTG	312	
Db	146	GAGGATAAATTGTGGCTTCCCGGTGTACCGCCAGCAGTGCACGGAGAGAGTTGCTG	205	
QY	313	TTTTGATGACAGTGTCCGGGATTCCTGGTGTCTCCACCCCATGGCCATCGAGAACAC	372	
Db	206	TTTTGATGACAGTGTCCGGGATTCCTGGTGTCTCCACCCCATGGCCATCGAGAACAC	265	
QY	373	TCAAGAAGAAGATGTCCCTTCTAA	397	
Db	266	TCAAGAAGAAGATGTCCCTTCTAA	290	

Search completed: September 24, 2004, 09:58:24  
Job time : 7806.56 secs



This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 18:06:49 ; Search time 219.773 Seconds  
(without alignments)  
12984.105 Million cell updates/sec

Title: US-10-030-390-1

Perfect score: 5142

Sequence: 1 gaattcgagctcggtaccg.....ccgtctgacgtctcttaag 5142

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	2467.2	48.0	8119	1 US-08-460-343B-1	Sequence 1, Appli
C 2	2467.2	48.0	8119	1 US-08-398-028B-1	Sequence 1, Appli
C 3	2467.2	48.0	8119	2 US-08-504-265B-1	Sequence 1, Appli
C 4	1636.8	31.8	3792	2 US-08-992-334-1	Sequence 1, Appli
C 5	1636.8	31.8	3792	3 US-08-302-752-1	Sequence 1, Appli
C 6	1635.2	31.8	5234	2 US-08-992-334-2	Sequence 2, Appli
C 7	1635.2	31.8	5234	3 US-08-302-752-2	Sequence 2, Appli
C 8	1635.2	31.8	6722	2 US-08-992-334-3	Sequence 2, Appli
C 9	1635.2	31.8	6722	3 US-08-302-752-3	Sequence 3, Appli
C 10	1493	29.0	7336	3 US-08-418-085-3	Sequence 3, Appli
C 11	1493	29.0	7336	3 US-09-099-011A-3	Sequence 3, Appli
C 12	1493	29.0	7336	3 US-08-470-368-21	Sequence 21, Appli
C 13	1493	29.0	7336	4 US-09-098-877B-3	Sequence 3, Appli
C 14	1493	29.0	9143	3 US-08-556-978B-79	Sequence 79, Appli
C 15	1491.4	29.0	7326	6 5304637-21	Patent No. 5304637
C 16	1479	28.8	5609	4 US-09-313-677-14	Sequence 14, Appl
C 17	1479	28.8	5609	4 US-09-313-677-15	Sequence 15, Appl
C 18	1479	28.8	7026	4 US-09-313-677-20	Sequence 20, Appl
C 19	1479	28.8	7344	4 US-09-313-677-16	Sequence 16, Appl
C 20	1084.6	21.1	6169	2 US-08-875-154-2	Sequence 2, Appli
C 21	867	16.9	3156	4 US-09-633-927-1	Sequence 1, Appli
C 22	830.8	16.2	10216	2 US-08-875-154-1	Sequence 1, Appli
C 23	763.8	14.9	1200	2 US-08-642-045B-3	Sequence 3, Appli
C 24	763.8	14.9	1200	3 US-08-852-268-1	Sequence 1, Appli
C 25	667.8	13.0	1402	1 US-08-297-294A-2	Sequence 2, Appli
C 26	408	7.9	3792	2 US-08-992-334-1	Sequence 1, Appli
C 27	408	7.9	3792	3 US-08-302-752-1	Sequence 1, Appli

RESULT 1  
US-08-460-343B-1/c  
; Sequence 1, Application US/08460343B  
; Patent No. 5741664  
; GENERAL INFORMATION:  
; APPLICANT: Marcus D. Ballinger and James A. Wells  
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING  
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA: US/08/460,343B  
; APPLICATION NUMBER: US/08/460,343B  
; FILING DATE: 01-Jun-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/398028  
; FILING DATE: 03-Mar-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rubinec, Jeffrey S.  
; REGISTRATION NUMBER: 36,575  
; REFERENCE/DOCKET NUMBER: P0936C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-8228  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8119 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
US-08-460-343B-1

ALIGNMENTS

C 28	408	7.9	5234	2	US-08-992-334-2	Sequence 2, Appli
C 29	408	7.9	5234	3	US-08-302-752-2	Sequence 2, Appli
C 30	408	7.9	6722	2	US-08-992-334-3	Sequence 3, Appli
C 31	408	7.9	6722	3	US-08-302-752-3	Sequence 3, Appli
C 32	347.2	6.8	5390	5	PCT-US96-12545-5	Sequence 5, Appli
C 33	162	3.2	5502	4	US-09-702-705-785	Sequence 785, App
C 34	162	3.2	5502	4	US-09-736-457-785	Sequence 785, App
C 35	162	3.2	5502	4	US-09-614-124B-785	Sequence 785, App
C 36	162	3.2	5502	4	US-09-671-325-785	Sequence 785, App
C 37	162	3.2	5502	4	US-09-589-184-785	Sequence 785, App
C 38	151.8	3.0	5217	4	US-09-100-703A-25	Sequence 25, Appli
C 39	150.8	2.9	5230	4	US-09-100-703A-26	Sequence 26, Appli
C 40	148.4	2.9	2321	3	US-09-198-839-2	Sequence 2, Appli
C 41	148.4	2.9	2321	4	US-09-562-834-2	Sequence 2, Appli
C 42	148.4	2.9	2321	4	US-09-995-396-2	Sequence 3, Appli
C 43	148.4	2.9	2337	3	US-09-198-839-3	Sequence 3, Appli
C 44	148.4	2.9	2337	4	US-09-562-834-3	Sequence 3, Appli
C 45	148.4	2.9	2337	4	US-09-995-396-3	Sequence 3, Appli

Query Match 48.0%; Score 2467.2; DB 1; Length 8119;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 831 TTGAAACCCGCTGTAAAGTACGTCGCGCATTCATTCATATATAAAGCCAGTCAT 890  
Db 7847 TTGAAACCCGCTGTAAAGTACGTCGCGCATTCATTCATATATAAAGCCAGTCAT 7788

QY 891 TAGGCTATCTGACAAATCCCTGAATAGAGTTTCATAAAACAATCCCTGCGATGATAACCATCAC 950  
Db 7787 TAGGCTATCTGACAAATCCCTGAATAGAGTTTCATAAAACAATCCCTGCGATGATAACCATCAC 7728  
QY 951 AAAACAGATGATGCTACCTGTAAGATAGCGGTAAATATATGAAATACCTTTATATATGA 1010  
Db 7727 AAAACAGATGATGCTACCTGTAAGATAGCGGTAAATATATGAAATACCTTTATATATGA 7668  
QY 1011 ATTTTCCTGCTGTAATATGCTGTAAGAGTAAATCTACTATTTATTTATTTATGATATTTAAGTTAA 1070  
Db 7667 ATTTTCCTGCTGTAATATGCTGTAAGAGTAAATCTACTATTTATTTATTTATGATATTTAAGTTAA 7608  
QY 1071 ACCCAGTAATGAAGTCAATGGAATATAGAAAGAGAAAAGCAATTTTCAGGTATAGGTG 1130  
Db 7607 ACCCAGTAATGAAGTCAATGGAATATAGAAAGAGAAAAGCAATTTTCAGGTATAGGTG 7548  
QY 1131 TTTTCGGGAAACAAATTTCCCGAACCAATATATTTCTCTACATCAGAAAGGTATATAATCAT 1190  
Db 7547 TTTTCGGGAAACAAATTTCCCGAACCAATATATTTCTCTACATCAGAAAGGTATATAATCAT 7488  
QY 1191 AAAACCTTTGAGTCAATCTTTTACAGGAGTCCAAATACAGAGAAATGTTTTAGATACAC 1250  
Db 7487 AAAACCTTTGAGTCAATCTTTTACAGGAGTCCAAATACAGAGAAATGTTTTAGATACAC 7428  
QY 1251 CATCAAAAATGTAATAAGTGGCTCTAACTTATCCCAATACCTAACTCTCCGTCGCTAT 1310  
Db 7427 CATCAAAAATGTAATAAGTGGCTCTAACTTATCCCAATACCTAACTCTCCGTCGCTAT 7368  
QY 1311 TGTAACCAAGTCTCTAAAAGCTGTAATTTGAGTTTATCACCCCTTGTCACTAAGAAAATAAATG 1370  
Db 7367 TGTAACCAAGTCTCTAAAAGCTGTAATTTGAGTTTATCACCCCTTGTCACTAAGAAAATAAATG 7308  
QY 1371 CAGGTAATAATTTATATCTCTGTTTATGTTTTCGGTATAAACAACCTAATATCAATTT 1430  
Db 7307 CAGGTAATAATTTATATCTCTGTTTATGTTTTCGGTATAAACAACCTAATATCAATTT 7248  
QY 1431 CTGTGGTTTATCTAAAAGTCTGTTTGGTTCAATAATGATTAATAATATCTCTTTCTCT 1490  
Db 7247 CTGTGGTTTATCTAAAAGTCTGTTTGGTTCAATAATGATTAATAATATCTCTTTCTCT 7188  
QY 1491 TCCAAATGCTAAATCAATTTTATTAAGTTCAATTTGATAGCTCTCCCTAAATTTTATCT 1550  
Db 7187 TCCAAATGCTAAATCAATTTTATTAAGTTCAATTTGATAGCTCTCCCTAAATTTTATCT 7128  
QY 1551 AAAGTGAATTTAGGAGGCTTACTCTGCTCTCTCTCAATAGAAATCAATCTTTTAAA 1610  
Db 7127 AAAGTGAATTTAGGAGGCTTACTCTGCTCTCTCTCAATAGAAATCAATCTTTTAAA 7068  
QY 1611 AGTCAATATCTGTAACATAAATAATATATTTTAAAATATCCCACTTTATCCAAATTTTC 1670  
Db 7067 AGTCAATATCTGTAACATAAATAATATATTTTAAAATATCCCACTTTATCCAAATTTTC 7008  
QY 1671 GTTCTGTAATTAAGTGGTCTTTAGTTGAAGATTAAGACCAATTAATAAATATGTCGTC 1730  
Db 7007 GTTCTGTAATTAAGTGGTCTTTAGTTGAAGATTAAGACCAATTAATAAATATGTCGTC 6948  
QY 1731 TTTTGTGTTTTTTTAAAGGATTTAGCGTAGCGAAAAATCTCTTTTCTTTTATCTTTGAT 1790  
Db 6947 TTTTGTGTTTTTTTAAAGGATTTAGCGTAGCGAAAAATCTCTTTTCTTTTATCTTTGAT 6888  
QY 1791 AATAAGGTAACTATGCGGATAGACTGTAACATTTCTCAGCATAAATCCCTTTCA 1850  
Db 6887 AATAAGGTAACTATGCGGATAGACTGTAACATTTCTCAGCATAAATCCCTTTCA 6828  
QY 1851 TTTTCTAATGTAATCTATTTACTTTATTAATTAATTTCAATTCGCTCATAAATTAATCCCTTTT 1910  
Db 6827 TTTTCTAATGTAATCTATTTACTTTATTAATTAATTTCAATTCGCTCATAAATTAATCCCTTTT 6768  
QY 1911 TCTTATACGAAAATGGCCGATTTAAGACACACCTTTTATTCGTTAATGCGGCATGAC 1970  
Db 6767 TCTTATACGAAAATGGCCGATTTAAGACACACCTTTTATTCGTTAATGCGGCATGAC 6708

QY 1971 AGCCATGATAATTAATACTAGGAGAGTTTAAATACTAGTACCAACATGATTACACA 2030  
Db 6707 AGCCATGATAATTAATACTAGGAGAGTTTAAATACTAGTACCAACATGATTACACA 6648  
QY 2031 ATTATTAGAGGTCAATCGTTTCAAAATGGTATGCTTTTGACACATCCCATATATATCCGTTG 2090  
Db 6647 ATTATTAGAGGTCAATCGTTTCAAAATGGTATGCTTTTGACACATCCCATATATATCCGTTG 6588  
QY 2091 TCGTTCTGTCCTCTGTAATCCCAATTCAGAAATTTCTCTAGCGATTCAGAGATTTCTC 2150  
Db 6587 TCGTTCTGTCCTCTGTAATCCCAATTCAGAAATTTCTCTAGCGATTCAGAGATTTCTC 6528  
QY 2151 AGAGTCGGAAGTTGACACAGACATTTACGAATCTGGCAGACAGATGGTCATACCTGAAGGAAG 2210  
Db 6527 AGAGTCGGAAGTTGACACAGACATTTACGAATCTGGCAGACAGATGGTCATACCTGAAGGAAG 6468  
QY 2211 ATCTGATTGCTTAACTGCTTCAAGTTAAGACCGAAGGCTCGTCGTATTAACAGATGCGATG 2270  
Db 6467 ATCTGATTGCTTAACTGCTTCAAGTTAAGACCGAAGGCTCGTCGTATTAACAGATGCGATG 6408  
QY 2271 ATGCGAGCCAAATCAACATGGCACCTGCTACCTGTACAGTCAAGGATGGTAGAAA 2330  
Db 6407 ATGCGAGCCAAATCAACATGGCACCTGCTACCTGTGTACAGTCAAGGATGGTAGAAA 6348  
QY 2331 TGTGTGCGTCTCTGACACAGCAATATTACGCCATTTGCTCGATATTCAACAGCTCTTC 2390  
Db 6347 TGTGTGCGTCTCTGACACAGCAATATTACGCCATTTGCTCGATATTCAACAGCTCTTC 6288  
QY 2391 TAGGATAAGGACAAATCGCATCTGGAAAGTTTGGCTTCTACCGATTTAGCAGTTTG 2450  
Db 6287 TAGGATAAGGACAAATCGCATCTGGAAAGTTTGGCTTCTACCGATTTAGCAGTTTG 6228  
QY 2451 ATACACTTTTCTCTAAGTATCCACCTGAATCATATAATCGGCAAAATAGAGAAAAATTGACC 2510  
Db 6227 ATACACTTTTCTCTAAGTATCCACCTGAATCATATAATCGGCAAAATAGAGAAAAATTGACC 6168  
QY 2511 ATGTGTAGCGGCGAATCTGATTCACCTGAGATGCAATACTAGTAGAATCTCTTCGCT 2570  
Db 6167 ATGTGTAGCGGCGAATCTGATTCACCTGAGATGCAATACTAGTAGAATCTCTTCGCT 6108  
QY 2571 ATCAAAATTCATCTCACCTTCCACTCCAGGTTGCTCAATTCATGGCTGAATCTGCTTC 2630  
Db 6107 ATCAAAATTCATCTCACCTTCCACTCCAGGTTGCTCAATTCATGGCTGAATCTGCTTC 6048  
QY 2631 CTCTGTGATGACACACATCATCTCAATATCCGAATAGGCGCCCATCAGTCTGACGACC 2690  
Db 6047 CTCTGTGATGACACACATCATCTCAATATCCGAATAGGCGCCCATCAGTCTGACGACC 5988  
QY 2691 AAGAGGCCATAAACAACAATAGCCTTAACATCATCCCATATTTATCCAAATATTCGTTTC 2750  
Db 5987 AAGAGGCCATAAACAACAATAGCCTTAACATCATCCCATATTTATCCAAATATTCGTTTC 5928  
QY 2751 CTTAATTTTCATGAACAAATCTTCAATCTTTCTCTCTAGTCAATTAATTTGGTCCATTCAC 2810  
Db 5927 CTTAATTTTCATGAACAAATCTTCAATCTTTCTCTCTAGTCAATTAATTTGGTCCATTCAC 5868  
QY 2811 TATTCTCATCTCCCTTTGAGATAATTTAGATTTGCTTTTCTAAATAAGAAATATTTGGAG 2870  
Db 5867 TATTCTCATCTCCCTTTGAGATAATTTAGATTTGCTTTTCTAAATAAGAAATATTTGGAG 5808  
QY 2871 AGCACCGTTCTTATTGAGCTATTAATACTCGTCTTCTCTTAAGCATCTTCAATCTTTTAA 2930  
Db 5807 AGCACCGTTCTTATTGAGCTATTAATACTCGTCTTCTCTTAAGCATCTTCAATCTTTTAA 5748  
QY 2931 ATAACAATTAAGCATCTTAATCTTCAACAACTGGCCGCTTTGTTGAATCTCTTTAAT 2990  
Db 5747 ATAACAATTAAGCATCTTAATCTTCAACAACTGGCCGCTTTGTTGAATCTCTTTAAT 5688  
QY 2991 AAAATAATTTTTCGTTTCCCAATTCACATTCGAATAATAGAAAAATCCATCTTCATCGGC 3050  
Db 5687 AAAATAATTTTTCGTTTCCCAATTCACATTCGAATAATAGAAAAATCCATCTTCATCGGC 5628  
QY 3051 TTTTTCGTCATCATCTGTATGAATCAAAATCGCCTTCTCTCTGTGTCATCAAGGTTTAAATTT 3110

Db 5627 TTTTCGTCATCATCTGTATGAATCAATCGCTTCTCTGTGTCATCAAGGTTTAATTT 5568  
Qy 3111 TTTATGTAATTTCTTTTACAAACACCATAGAGATTAAACCTTTTACGGTGTAAACCTTC 3170  
Db 5567 TTTATGTAATTTCTTTTACAAACACCATAGAGATTAAACCTTTTACGGTGTAAACCTTC 5508  
Qy 3171 CTCCAAATCAGACAAACGTTTCAAAATCTTTCTTCATCATCGGTCAATAAAATCCGTATC 3230  
Db 5507 CTCCAAATCAGACAAACGTTTCAAAATCTTTCTTCATCATCGGTCAATAAAATCCGTATC 5448  
Qy 3231 CTTTACAGATATTTTTCAGATTCTGCAATTCGCGATTGTATCCGATTATATTTATT 3290  
Db 5447 CTTTACAGATATTTTTCAGATTCTGCAATTCGCGATTGTATCCGATTATATTTATT 5388  
Qy 3291 TTTTCGGTATTTTATTATAA 3310  
Db 5387 TTTTCGGTCAATCATTTGAA 5368

RESULT 2  
US-08-398-028B-1/c  
; Sequence 1, Application US/08398028B  
; Patent No. 5780285  
; GENERAL INFORMATION:  
; APPLICANT: Marcus D. Ballinger and James A. Wells  
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING  
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Winpatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/398,028B  
; FILING DATE: 03-Mar-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kubinec, Jeffrey S.  
; REGISTRATION NUMBER: 36,575  
; REFERENCE/DOCKET NUMBER: P0936  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-8228  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8119 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; US-08-398-028B-1

Query Match 48.0%; Score 2467.2; DB 1; Length 8119;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 831 TTGAGAAACCGAGCTGTAAAGATACAGTCGGCATTTATCTCATATTAATAAAGCCAGTCAT 890  
Db 7847 TTGAGAAACCGAGCTGTAAAGATACAGTCGGCATTTATCTCATATTAATAAAGCCAGTCAT 7788  
Qy 891 TAGGCCTATCTGCAATTTCTCGAATAGATTCAATAAACAATCTCTGATGATCAATCAATCAC 950  
Db 7787 TAGGCCTATCTGCAATTTCTCGAATAGATTCAATAAACAATCTCTGATGATCAATCAATCAC 7728  
Qy 951 AAACAGATGATCTACTGTAAAGATAGCGGTAAATATATTGAAATTAACCTTTTATTATGCA 1010

Db 7727 AAACAGATGATCTACTGTAAAGATAGCGGTAAATATATTGAAATTAACCTTTATTATGCA 7668  
Qy 1011 ATTTTCTCTGTATTAATCGGTAGAGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1070  
Db 7667 ATTTTCTCTGTATTAATCGGTAGAGGTAATTAATTAATTAATTAATTAATTAATTAATTA 7608  
Qy 1071 ACCCAGTAATGAAGTCCATGTAATTAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1130  
Db 7607 ACCCAGTAATGAAGTCCATGTAATTAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7548  
Qy 1131 TTTTGGGAAACAATTTCCCGGAACCAATTAATTTCTCTACATCAGAGAGAGAGAGAGAGAGAGAG 1190  
Db 7547 TTTTGGGAAACAATTTCCCGGAACCAATTAATTTCTCTACATCAGAGAGAGAGAGAGAGAG 7488  
Qy 1191 AAAACTCTTTGAAGTCAATTTTACAGGAGTCCAAATACAGAGAGAGAGAGAGAGAGAGAGAG 1250  
Db 7487 AAAACTCTTTGAAGTCAATTTTACAGGAGTCCAAATACAGAGAGAGAGAGAGAGAGAGAG 7428  
Qy 1251 CATCAAAAATTTGTATAAAGTGGCTTAACTTATCCCAATAAACCCTAACTCTCCGTCGCTAT 1310  
Db 7427 CATCAAAAATTTGTATAAAGTGGCTTAACTTATCCCAATAAACCCTAACTCTCCGTCGCTAT 7368  
Qy 1311 TGTAACCAAGTCTAAAAGGCTGTATTTGAGTTATCAACCCCTTGTCACCTAGAGAGAGAGAGAG 1370  
Db 7367 TGTAACCAAGTCTAAAAGGCTGTATTTGAGTTATCAACCCCTTGTCACCTAGAGAGAGAGAG 7308  
Qy 1371 CAGGTTAAAATTTATATCTCTCTTTGTTATGTTGCGGTATTAATAAACAATTAATTAATTAAT 1430  
Db 7307 CAGGTTAAAATTTATATCTCTCTTTGTTATGTTGCGGTATTAATAAACAATTAATTAATTAAT 7248  
Qy 1431 CTGTGTTATATAAAGTCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1490  
Db 7247 CTGTGTTATATAAAGTCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 7188  
Qy 1491 TCCAAATGCTAAATCAATTTTATTAAGTTTCAATTTGATATGCTCTCTAAATTTTATCT 1550  
Db 7187 TCCAAATGCTAAATCAATTTTATTAAGTTTCAATTTGATATGCTCTCTAAATTTTATCT 7128  
Qy 1551 AAAGTGAATTTAGGAGGCTTACTGTCTGCTTTCTTCATTAGAGATCAATCTTTTAAAT 1610  
Db 7127 AAAGTGAATTTAGGAGGCTTACTGTCTGCTTTCTTCATTAGAGATCAATCTTTTAAAT 7068  
Qy 1611 AGTCAATATTAAGTGAATCAATTAATTTTAAATAATCCCACTTTATCAATTTTTC 1670  
Db 7067 AGTCAATATTAAGTGAATCAATTAATTTTAAATAATCCCACTTTATCAATTTTTC 7008  
Qy 1671 GTTTGTTGAATTAAGTGGTCTTTAGTTGAAGATTAAGACCAATTAATAAATTTGCTGTC 1730  
Db 7007 GTTTGTTGAATTAAGTGGTCTTTAGTTGAAGATTAAGACCAATTAATAAATTTGCTGTC 6948  
Qy 1731 TTTTGTGTTTTTTTAAAGGATTTGAGCGTAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1790  
Db 6947 TTTTGTGTTTTTTTAAAGGATTTGAGCGTAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6888  
Qy 1791 AATAAGGATTAATTTCCCGGATAGACTGTACATTTCTCAGCATATAAATCCCTTTTCA 1850  
Db 6887 AATAAGGATTAATTTCCCGGATAGACTGTACATTTCTCAGCATATAAATCCCTTTTCA 6828  
Qy 1851 TTTTCTAATGTAAATCTATTACCTTTATTTAATTTCAATTCGGTCAATAAATTAATTCCTTTT 1910  
Db 6827 TTTTCTAATGTAAATCTATTACCTTTATTTAATTTCAATTCGGTCAATAAATTAATTCCTTTT 6768  
Qy 1911 TCTTATTACGAAATAGCGCGGATTTAAGCACACCTTTTATTCGGTTAATTCGGCGCATGAC 1970  
Db 6767 TCTTATTACGAAATAGCGCGGATTTAAGCACACCTTTTATTCGGTTAATTCGGCGCATGAC 6708  
Qy 1971 AGCCATGATATTAATTAATTAAG 2030  
Db 6707 AGCCATGATATTAATTAATTAAG 6648  
Qy 2031 ATTATTAGAGTCTGTTTCAAAATGGTATCGGTTTTTGACACATCCCATATATATTCGGTG 2090





Db 5507 CTCCAATCAGACAAACGTTTCAAATCTTTTCTTCATCGCGTCATAAAATCCGATC 5448  
Qy 3231 CTTTACAGATATTTGAGATTTGCTCAATGCGGATGTATATCCGATTTATATTTAT 3290  
Db 5447 CTTTACAGATATTTGAGATTTGCTCAATGCGGATGTATATCCGATTTATATTTAT 5388  
Qy 3291 TTTTCGGTATTTTATTTAAA 3310  
Db 5387 TTTTCGGTCAATCATTTGAA 5368

RESULT 4  
US-08-992-334-1  
; Sequence 1, Application US/08992334  
; Patent No. 5919678  
; GENERAL INFORMATION:  
; APPLICANT: Gruss, Alexandra  
; APPLICANT: Maguin, Emmanuelle  
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE  
; TITLE OF INVENTION: PLASMID  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Christie Parker & Hale, LLP  
; STREET: 350 West Colorado Boulevard, Suite 500  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: United States  
; ZIP: 91105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/992,334  
; FILING DATE: 17-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/302,752  
; FILING DATE: 24-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR93/00248  
; FILING DATE: 12-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 992/03034  
; FILING DATE: 13-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Prout, D. Bruce  
; REGISTRATION NUMBER: 20958  
; REFERENCE/DOCKET NUMBER: C93:31779  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (626) 795-9900  
; TELEFAX: (626) 577-8800  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3792 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: YES  
; IMMEDIATE SOURCE:  
; CLONE: pg+host4  
; US-08-992-334-1

Query Match 31.8%; Score 1636.8; DB 2; Length 3792;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 1686; Conservative 0; Mismatches 27; Indels 3; Gaps 3;

Qy 3279 TTTATATTTATTTTTCGGTATTTTATTTAAACGCTCTCAAAATCGTTTCTCGGACGTTT 3338  
Db 1181 TATATATTTATTTATTCGCAATTTTATTTAAACGCTCTCAAAATCGTTTCTCGGACGTTT 1240

Qy 3339 TAGCGTTTATTTTCGTTTAGTTATCGGCATAATCGTTAAAAACAGCGTTTATCGTAGCGTAA 3398  
Db 1241 TAGCGTTTATTTTCGTTTAGTTATCGGCATAATCGTTAAAAACAGCGTTTATCGTAGCGTAA 1300  
Qy 3399 AAGCCCTTGAGCGTAGCGT-GGTTTGCAGCGAAGATGTTGTTCTGTATGATTATGAAGGCC 3457  
Db 1301 AAGCCCTTGAGCGTAGCGTGGCTTTGCGAGCGAAGATGTTGTTCTGTATGATTATGAAGGCC 1360  
Qy 3458 GATGACTGAATGAAATAATAAGCGCAGCGTCCCTCTATTTCCGTTGGAGGAGGCTCAAGG 3517  
Db 1361 GATGACTGAATGAAATAATAAGCGCAGCGCCCTCTATTTCCGTTGGAGGAGGCTCAAGG 1420  
Qy 3518 GAGTTTGAGGGAATGAAATTCCTCATCGGTTTGATTTTAAAAATGCTTCGCAATTTTGC 3577  
Db 1421 GAGTATGAGGGAATGAAATTCCTCATCGGTTTGATTTTAAAAATGCTTCGCAATTTTGC 1480  
Qy 3578 CGAGCGGTAGCGCTCGAAATTTTGAATAAATTTGGAATTTGGAATAAATGGAATAAATGGGGGA 3637  
Db 1481 CGAGCGGTAGCGCTCGAAATTTTGAATAAATTTTGAATAAATTTGGAATAAATGGGGGA 1540  
Qy 3638 AAGGAAGCGAATTTTTCGTTCCGTACTAGACCCCGCATTAAGTCGCGAGTGCCTATTTT 3697  
Db 1541 AAGGAAGCGAATTTTTCGTTCCGTACTAGACCCCGCATTAAGTCGCGAGTGCCTATTTT 1600  
Qy 3698 GTGCCAAAACGCTCTATCCCACTGGCTCAAGGGTTTGAAGGGTTTTCATTCGCCAAC 3757  
Db 1601 GTGCCAAAACGCTCTATCCCACTGGCTCAAGGGTTTGAAGGGTTTTCATTCGCCAAC 1660  
Qy 3758 GAATCGCCAAACGTTTTCGCCAACGTTTTCGCAACGTTTTCGCAACGTTTTCGCAACGTTTTCGTT 3817  
Db 1661 GAATCGCCAAACGTTTTCGCCAACGTTTTCGCAACGTTTTCGCAACGTTTTCGCAACGTTTTCGTT 1720  
Qy 3818 GTTTTATGATTAACAAGTATACACTATTTTATAAATAATTTTGAATTTGAGTTTGAATTTTGA 3877  
Db 1721 GTTTTATGATTAACAAGTATACACTATTTTATAAATAATTTTGAATTTTGAATTTTGAATTTTGA 1780  
Qy 3878 AATGTTGATTTTCAGAAATCGAAAAAAGAGTTTATGATTTCTCTGCAAAAAGAGACAGATAA 3937  
Db 1781 AATGTTGATTTTCAGAAATCGAAAAAAGAGTTTATGATTTCTCTGCAAAAAGAGACAGATAA 1840  
Qy 3938 AAAATTAAACAGATATGCGGAAACAAAGAGTTTTCGCAATCTGCGTTGCGGCGTTAGC 3997  
Db 1841 AAAATTAAACAGATATGCGGAAACAAAGAGTTTTCGCAATCTGCGTTGCGGCGTTAGC 1900  
Qy 3998 TATAGAAGAAATATGCAAGAAAGGAATCAGAAACAAAAAATAAGCGAAAGCTCGCGTTT 4057  
Db 1901 TATAGAAGAAATATGCAAGAAAGGAATCAGAAACAAAAAATAAGCGAAAGCTCGCGTTT 1960  
Qy 4058 TAGAAGGATACGAGTTTTCGTTACTGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 4116  
Db 1961 TAGAAGGATACGAGTTTTCGTTACTGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 2020  
Qy 4117 AATATCTAAAGCTAGAAATTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 4176  
Db 2021 AATATCTAAAGCTAGAAATTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 2080  
Qy 4177 GAAAGAAAAATTAGAGAGTTTGGCGGTATCTATGCGTGTGCTGCTTTCACGATATGA 4236  
Db 2081 GAAAGAAAAATTAGAGAGTTTGGCGGTATCTATGCGTGTGCTGCTTTCACGATATGA 2140  
Qy 4237 CGAAAAAAGATTAAGATACATGGAATAGTGTGATTTATACGAAATGGAAGACCTA 4296  
Db 2141 CGAAAAAAGATTAAGATACATGGAATAGTGTGATTTATACAAAAATGGAAGACCTA 2200  
Qy 4297 TAAAAAACCACTATCAGCTTATATATATGCGAAATCTGTGTAAACATGAAGAGCGT 4356  
Db 2201 TAAAAAACCACTATCAGCTTATATATATGCGAAATCTGTGTAAACATGAAGAGCGT 2260  
Qy 4357 TAGGAACAAAGTTAAGCGAAATTTGGGATAGTTCAGTTGCTCATGTTGAGATATCTGA 4416  
Db 2261 TAGGAACAAAGTTAAGCGAAATTTGGGATAGTTCAGTTGCTCATGTTGAGATATCTGA 2320



QY 4417 TTATATCAAGGTTTCATATGAATATTTGACTCATGAATCAAGGACGCTATTGCTAAGAA 4476  
Db 2321 TTATATCAAGGTTTCATATGAATATTTGACTCATGAATCAAGGACGCTATTGCTAAGAA 2380  
QY 4477 TAAACATATATACGCAAAAGAGATATTTGAAACATTAATGATTTTGATATTTGACCGCTA 4536  
Db 2381 TAAACATATATACGCAAAAGAGATATTTGAAACATTAATGATTTTGATATTTGACCGCTA 2440  
QY 4537 TATAACACTTGATGAAGCAAAAGAGATTTGAAAGATTTGATCTTTTAGATATAGTGGG 4596  
Db 2441 TATAACACTTGATGAAGCAAAAGAGATTTGAAAGATTTGATCTTTTAGATATAGTGGG 2500  
QY 4597 TGACTATATTTGGTAAATACAAAGATTTAATGCTTTTATCGCCCTTAGCGGAGCGGA 4656  
Db 2501 TGACTATATTTGGTAAATACAAAGATTTAATGCTTTTATCGCCCTTAGCGGAGCGGA 2560  
QY 4657 GTTTCGAATTTTAAATACGAATGATGATAAAGATATTTGTTTCAACAAACTCTAGCGCCTT 4716  
Db 2561 GTTTCGAATTTTAAATACGAATGATGATAAAGATATTTGTTTCAACAAACTCTAGCGCCTT 2620  
QY 4717 TAGATATGTTTGGGCAATTTATCAGTGTGATATAGCAAGTTATGCAAGGTTCT 4776  
Db 2621 TAGATATGTTTGGGCAATTTATCAGTGTGATATAGCAAGTTATGCAAGGTTCT 2680  
QY 4777 TGATGCTGAAACGGGGAAATTAATGACAAACAAAGAAAGAGTATTTGCTGAAAT 4836  
Db 2681 TGATGCTGAAACGGGGAAATTAATGACAAACAAAGAAAGAGTATTTGCTGAAAT 2740  
QY 4837 GAGGAATTAATAAGAAATTAAGGACTTAAAGAGCGTATTTGAAAGATACAGAGAAATG 4896  
Db 2741 GAGGAATTAATAAGAAATTAAGGACTTAAAGAGCGTATTTGAAAGATACAGAGAAATG 2800  
QY 4897 GAAGTTGAATTAAGTACAAATAGATTTATTGAGAGGAGGATTTATTGAATAAATAAA 4956  
Db 2801 GAAGTTGAATTAAGTACAAATAGATTTATTGAGAGGAGGATTTATTGAATAAATAAA 2860  
QY 4957 GCGCCCTGACAAAGTCCGACTTCGTTCTTTTTT 4992  
Db 2861 G-CCGCTGACAAAGTCCGAGGGGTTTATTATTT 2895

## RESULT 5

US-08-302-752-1  
; Sequence 1, Application US/08302752  
; Patent No. 6025190  
; GENERAL INFORMATION:  
; APPLICANT: THERMOSENSIBLE PLASMID  
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID  
; NUMBER OF SEQUENCES: 3  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/302,752  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9203034  
; FILING DATE: 13-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO FR/93/00248  
; FILING DATE: 12-MAR-1993  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3792 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-302-752-1

Query Match 31.9%; Score 1636.8; DB 3; Length 3792;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 1686; Conservative 0; Mismatches 27; Indels 3; Gaps 3;  
QY 3279 TTTATATTTATTTTTCGGTATATTTTATTAATAAACGCTCATAAATCGTTCTCGGACGTTT 3338  
Db 1181 TATATATATTTATATCGCATTTTATTAATAACGCTCATAAATCGTTCTCGGACGTTT 1240  
QY 3339 TAGCGTTTATTTTCGTTAGTATTCGGCATTAATCGTTTAAACAGCGGCTTATCGTAGCGTAA 3398  
Db 1241 TAGCGTTTATTTTCGTTAGTATTCGGCATTAATCGTTTAAACAGCGGCTTATCGTAGCGTAA 1300  
QY 3399 AAGCCCTTGAGCGTACGCT-GCTTTCAGCGAGAGATGTTGCTGTTAGATTTATGAAGCC 3457  
Db 1301 AAGCCCTTGAGCGTACGCTGCTTTCAGCGAGAGATGTTGCTGTTAGATTTATGAAGCC 1360  
QY 3458 GATGACTGAATGAATAAATAAGCGCAGCGCTCTTCTTATTTTCGTTTCGAGAGGCTCAAGG 3517  
Db 1361 GATGACTGAATGAATAAATAAGCGCAGCGCTCTTCTTATTTTCGTTTCGAGAGGCTCAAGG 1420  
QY 3518 GAGTTTTCAGGGAATGAATTCCTCATGCTGTTGATTTTAAATAATGCTTGCATTTTTC 3577  
Db 1421 GAGTATGAGGAATGAATTCCTCATGCTGTTGATTTTAAATAATGCTTGCATTTTTC 1480  
QY 3578 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTGGAATTTGGAATAAATTTGCGGGGA 3637  
Db 1481 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTGGAATTTGGAATAAATTTGCGGGGA 1540  
QY 3638 AAGAGCGGAATTTTCGTTCCGCTACTACGACCCCGCATTAAGTCCGAGTGCCAAATTTT 3697  
Db 1541 AAGAGCGGAATTTTCGTTCCGCTACTACGACCCCGCATTAAGTCCGAGTGCCAAATTTT 1600  
QY 3698 GTGCCAAAACGCTCTATCCCACTGGCTCAAGGGTTTGAGGGTTTTCATTCGCGCAAC 3757  
Db 1601 GTGCCAAAACGCTCTATCCCACTGGCTCAAGGGTTTGAGGGTTTTCATTCGCGCAAC 1660  
QY 3758 GAATCGCCAAACGTTTTCGCAACGTTTTCATTAATCTATTAATTAAGTAGCTTTATGTT 3817  
Db 1661 GAATCGCCAAACGTTTTCGCAACGTTTTCATTAATCTATTAATTAAGTAGCTTTATGTT 1720  
QY 3818 GTTTTATGATTACAAAGTATACACTAATTTTATATAAATTTATGATGAGGTTTATTA 3877  
Db 1721 GTTTTATGATTACAAAGTATACACTAATTTTATATAAATTTATGATGAGGTTTATTA 1780  
QY 3878 AATGGTGATTTCAGAAATCGAAAAAAGAGTATGATTTCTCTGACAAAAAGAGCAAGATA 3937  
Db 1781 AATGGTGATTTCAGAAATCGAAAAAAGAGTATGATTTCTCTGACAAAAAGAGCAAGATA 1840  
QY 3938 AAAATTACAGATATGCGGCAACAAAGGTTTTCATAATCTCGCGTTGCGCGCTTAGC 3997  
Db 1841 AAAATTACAGATATGCGGCAACAAAGGTTTTCATAATCTCGCGTTGCGCGCTTAGC 1900  
QY 3998 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGGAAAGCTCGCGTTT 4057  
Db 1901 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGGAAAGCTCGCGTTT 1960  
QY 4058 TAGAAGATACAGGTTTTCGCTACTTGTGATTAAGGTAA-TATATCATGCTATTAA 4116  
Db 1961 TAGAAGATACAGGTTTTCGCTACTTGTGATTAAGGTAA-TATATCATGCTATTAA 2020  
QY 4117 AAATCTAAAGCTAGAAATTTTGGATTTTATTTATATCTGACTCAATTCCTAATGATG 4176  
Db 2021 AAATCTAAAGCTAGAAATTTTGGATTTTATTTATATCTGACTCAATTCCTAATGATG 2080  
QY 4177 GAAAGAAAAATTAGAGGTTTGGCGTATCTATGCTGTGCTCAGTCCCTTTACACGATAGGA 4236  
Db 2081 GAAAGAAAAATTAGAGGTTTGGCGTATCTATGCTGTGCTCAGTCCCTTTACACGATAGGA 2140  
QY 4237 CGAAAAAAGATAAAGATACATGGAATAGTAGTGTATATACGAATGGAAGCACTA 4296  
Db 2141 CGAAAAAAGATAAAGATACATGGAATAGTAGTGTATATACGAATGGAAGCACTA 2200  
QY 4297 TAAAAAACCACTATCAGCTTATATATATTTGACCAAGAAATCTCTGTAAACAATAGAAAGCGT 4356





QY 3818 GTTTTATGATTACAAAGTGATACACTAAATTTTATATAATTTATTTGATGGAGTTTTTTA 3877  
DB |||||  
QY 3163 GTTTTATGATTACAAAGTGATACACTAAATTTTATATAATTTATTTGATGGAGTTTTTTA 3222  
DB |||||  
QY 3878 AATGCTGATTTGAGAAATCGAAAAAGAGTTATGATTTCTCTGACAAAAGCAAGATAA 3937  
DB |||||  
QY 3223 AATGCTGATTTGAGAAATCGAAAAAGAGTTATGATTTCTCTGACAAAAGCAAGATAA 3282  
DB |||||  
QY 3938 AAAATTAACAGATATGCGGCAACCAAAAGAGTTTTCAAAATCTCGGTTGCGGCTTAGC 3997  
DB |||||  
QY 3283 AAAATTAACAGATATGCGGCAACCAAAAGAGTTTTCAAAATCTCGGTTGCGGCTTAGC 3342  
DB |||||  
QY 3998 TATAGAAGATATGCAAGAAAGGATCAGAACAAAAAATAGCGAAAGTCGCGTTTT 4057  
DB |||||  
QY 3343 TATAGAAGATATGCAAGAAAGGATCAGAACAAAAAATAGCGAAAGTCGCGTTTT 3402  
DB |||||  
QY 4058 TAGAAGGATACAGATTTTTCGCTACTTTTGTGATAGGTAA-TATATCATGGCTATTAA 4116  
DB |||||  
QY 3403 TAGAAGGATACAGATTTTTCGCTACTTTTGTGATAGGTAA-TATATCATGGCTATTAA 3462  
DB |||||  
QY 4117 AAATCTAAAGCTAGAAATTTTGGATTTTATATATCTCTGACTCAATTCCTATGATTG 4176  
DB |||||  
QY 3463 AAATCTAAAGCTAGAAATTTTGGATTTTATATATCTCTGACTCAATTCCTATGATTG 3522  
DB |||||  
QY 4177 GAAAGAAAATTAGAGAGTTTGGGCGTATCTATGCTGTCTAGTCTTTACACGATATGGA 4236  
DB |||||  
QY 3523 GAAAGAAAATTAGAGAGTTTGGGCGTATCTATGCTGTCTAGTCTTTACACGATATGGA 3582  
DB |||||  
QY 4237 CGAAAAAAGATTAAGATACATGATAGTATGATGATATATACGAAATGGAAGCACTA 4296  
DB |||||  
QY 3583 CGAAAAAAGATTAAGATACATGATAGTATGATGATATATACGAAATGGAAGCACTA 3642  
DB |||||  
QY 4297 TAAAAACCACTATACGTTATATATATGTCACGAAATCTGTAACATAGAAAGCGT 4356  
DB |||||  
QY 3643 TAAAAACCACTATACGTTATATATATGTCACGAAATCTGTAACATAGAAAGCGT 3702  
DB |||||  
QY 4357 TAGGACAGATTAAGGAAATTCGGGAATAGTTTCAGTGTCTATGTTGAGATCTTGA 4416  
DB |||||  
QY 3703 TAGGACAGATTAAGGAAATTCGGGAATAGTTTCAGTGTCTATGTTGAGATCTTGA 3762  
DB |||||  
QY 4417 TTATATCAAAGGTTTCATATGAATTTTGACTCATGAATCAAAGGACGCTATTGCTAGAA 4476  
DB |||||  
QY 3763 TTATATCAAAGGTTTCATATGAATTTTGACTCATGAATCAAAGGACGCTATTGCTAGAA 3822  
DB |||||  
QY 4477 TAAACATATACGACAAAAGATATTTTGAACATTAATGATTTTGTATTTGACCGCTA 4536  
DB |||||  
QY 3823 TAAACATATACGACAAAAGATATTTTGAACATTAATGATTTTGTATTTGACCGCTA 3882  
DB |||||  
QY 4537 TAAACATATGATGAAAGCCAAAAGAGATTTGAAGATTTTACTTTTAGATATAGTGA 4596  
DB |||||  
QY 3883 TAAACATATGATGAAAGCCAAAAGAGATTTGAAGATTTTACTTTTAGATATAGTGA 3942  
DB |||||  
QY 4597 TGACTATAATTTGCTGTHAATAACAAAGATTTATGCTTTTATTCGCTTAGGGAGCGGA 4656  
DB |||||  
QY 3943 TGACTATAATTTGCTGTHAATAACAAAGATTTATGCTTTTATTCGCTTAGGGAGCGGA 4002  
DB |||||  
QY 4657 GTTTGGAATTTTAAATACCAATGATGTAAGATATTTGTTCAACAAACTCTAGCGCCTT 4716  
DB |||||  
QY 4003 GTTTGGAATTTTAAATACCAATGATGTAAGATATTTGTTCAACAAACTCTAGCGCCTT 4062  
DB |||||  
QY 4717 TAGATTATGTTTGGGGCAATTTATCAGTGTGGATATAGAGCAAGTTATGCAAGGTTCT 4776  
DB |||||  
QY 4063 TAGATTATGTTTGGGGCAATTTATCAGTGTGGATATAGAGCAAGTTATGCAAGGTTCT 4122  
DB |||||  
QY 4777 TGATGCTGAACCGGGGAAATTAATATGCAACAAAGAAAGAGTTATTTCTCTGAAAT 4836  
DB |||||  
QY 4123 TGATGCTGAACCGGGGAAATTAATATGCAACAAAGAAAGAGTTATTTCTGTAAT 4182  
DB |||||  
QY 4837 GAGGAATTAATAAAGAAATTAAGGCTTTAAAGAGCGTATTGAAAGATACAGAAATG 4896  
DB |||||  
QY 4183 GAGGAATTAATAAAGAAATTAAGGCTTTAAAGAGCGTATTGAAAGATACAGAAATG 4242  
DB |||||  
QY 4897 GAAGTTGAATTAAGTACCAATAGATTTATTGAGAGGAGGATTTATTGAATAAATAAAA 4956  
DB |||||

DB 4243 GAAGTTGAATTAAGTACCAATAGATTTATTGAGAGGAGGATTTATTGAATAAATAAAA 4302  
QY 4957 GCCCCTCAGCAAGAGTCGCGACTTCGTTCTTTT 4992  
DB 4303 G-CCCCCTGACGAAAGTCGAAGGGGTTTATTATTT 4337  
RESULT 8  
US-08-992-334-3  
; Sequence 3, Application US/08992334  
; Patent No. 5919678  
; GENERAL INFORMATION:  
; APPLICANT: Gruss, Alexandra  
; APPLICANT: Maguin, Emmanuelle  
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE  
; TITLE OF INVENTION: PLASMID  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christie Parker & Hale, LLP  
; STREET: 350 West Colorado Boulevard, Suite 500  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: United States  
; ZIP: 91105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/992,334  
; FILING DATE: 17-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/302,752  
; FILING DATE: 24-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR93/00248  
; FILING DATE: 12-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 92/03034  
; FILING DATE: 13-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Prout, D. Bruce  
; REGISTRATION NUMBER: 20958  
; REFERENCE/DOCKET NUMBER: C93:31779  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (626) 795-9900  
; TELEFAX: (626) 577-8800  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6722 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
US-08-992-334-3  
Query Match 31.8%; Score 1635.2; DB 2; Length 6722;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 1685; Conservative 0; Mismatches 28; Indels 3; Gaps 3;  
QY 3279 TTATATATTTTTCGTTATTTTATTAACGCTCAAAATCGTTTCTGGAGCTTT 3338  
DB 4111 TATATATTTTATTTATTCGATTTTATTAACGCTCAAAATCGTTTCTGGAGCTTT 4170  
QY 3339 TAGGTTTATTTTCGTTATTTAGTTATCGGCATAATCGTTTAAACAGCGCTTATCGTAGCGTAA 3398  
DB 4171 TAGGTTTATTTTCGTTATTTAGTTATCGGCATAATCGTTTAAACAGCGCTTATCGTAGCGTAA 4230  
QY 3399 AAGCCCTTGAGCGTAGCGT-GCTTTGAGGAAGATGTTGTTCTGTAGATTATGAAGCC 3457  
DB |||||



3339 TAGCGTTTATTCGTTAGTTATCGGCAATATCGTTAAAAACAGCGGTTATCGTAGCGTAA 3398  
Db 4171 TAGCGTTTATTCGTTAGTTATCGGCAATATCGTTAAAAACAGCGGTTATCGTAGCGTAA 4230  
QY 3399 AAGCCCTTACGCGTAGCGT - CTTTTCAGCGAAGATGTTCTGTTAGATATTGAAAGCC 3457  
Db 4231 AAGCCCTTACGCGTAGCGT - CTTTTCAGCGAAGATGTTCTGTTAGATATTGAAAGCC 4290  
QY 3458 GATGACTGAATGAATAATATAGCGAGCGTCTCTTCTATTTTCGTTGAGGAGCGTCAAGG 3517  
Db 4291 GATGACTGAATGAATAATATAGCGAGCGTCTCTTCTATTTTCGTTGAGGAGCGTCAAGG 4350  
QY 3518 GAGTTTGAAGGAATGAATTCCTTCCTGCTGTTGATTTTAAATTCCTTGCCTATTTTCG 3577  
Db 4351 GAGTTTGAAGGAATGAATTCCTTCCTGCTGTTGATTTTAAATTCCTTGCCTATTTTCG 4410  
QY 3578 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTGGAATTTGGAATAAATTTGGAATAA 3637  
Db 4411 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTGGAATTTGGAATAAATTTGGAATAA 4470  
QY 3638 AAGGAGCGAATTTTTCCTGCTGCTAGCGACCCCGATTAAGTCCGAGTGCCTAATTTT 3697  
Db 4471 AAGGAGCGAATTTTTCCTGCTGCTAGCGACCCCGATTAAGTCCGAGTGCCTAATTTT 4530  
QY 3698 GTGCCAAAACGCTCTATCCCACTGCTCAAGGCTTTGAGGCGTCTTTCAATCGCCAAC 3757  
Db 4531 GTGCCAAAACGCTCTATCCCACTGCTCAAGGCTTTGAGGCGTCTTTCAATCGCCAAC 4590  
QY 3758 GAATCGCAACGTTTTCGCGCAACGTTTTCCTAATCTATTTAAGTAGCTTTATTTGTT 3817  
Db 4591 GAATCGCAACGTTTTCGCGCAACGTTTTCCTAATCTATTTAAGTAGCTTTATTTGTT 4650  
QY 3818 GTTTTATGATTACAAAGTGATACACTAAATTTTATAAATTTATTTGATTCGAGTTTTTA 3877  
Db 4651 GTTTTATGATTACAAAGTGATACACTAAATTTTATAAATTTATTTGATTCGAGTTTTTA 4710  
QY 3878 AATGCTGATTTTCAAGATCGAATAAAGAGTTATGATTTCTCTGACAAAGAGCAAGATAA 3937  
Db 4711 AATGCTGATTTTCAAGATCGAATAAAGAGTTATGATTTCTCTGACAAAGAGCAAGATAA 4770  
QY 3938 AAAATTAACAGATATGCGCAACAAAGAGTTTTCCTAATCTCGGTTGCGGCTTAGC 3997  
Db 4771 AAAATTAACAGATATGCGCAACAAAGAGTTTTCCTAATCTCGGTTGCGGCTTAGC 4830  
QY 3998 TATAGAAGATATGCAAGAAAGAGTATCAGAACAAATAAAGTAAAGTAAAGTAAAGTAA 4957  
Db 4831 TATAGAAGATATGCAAGAAAGAGTATCAGAACAAATAAAGTAAAGTAAAGTAAAGTAA 4890  
QY 4058 TAGAAGGATACAGTTTTCGCTACTGTTTGTGATAGGTA - TATATCATGCTATTAA 4116  
Db 4891 TAGAAGGATACAGTTTTCGCTACTGTTTGTGATAGGTA - TATATCATGCTATTAA 4950  
QY 4117 AAATACTAAAGCTAGAAATTTTGGATTTTATATCTGACTCAATTTCTAATGATG 4176  
Db 4951 AAATACTAAAGCTAGAAATTTTGGATTTTATATCTGACTCAATTTCTAATGATG 5010  
QY 4177 GAAAGAAAATTTAGAGGTTTGGCGTATCTATGCTGCTGCTGCTTACAGTATGGA 4236  
Db 5011 GAAAGAAAATTTAGAGGTTTGGCGTATCTATGCTGCTGCTTACAGTATGGA 5070  
QY 4237 CGAAAAAAGATTAAGATACATGGAATAGTAGTATGATGATGATGATGATGATGATGATG 4296  
Db 5071 CGAAAAAAGATTAAGATACATGGAATAGTAGTATGATGATGATGATGATGATGATGATG 5130  
QY 4297 TAAAAAACCACTATCATGTTATATATTTGCGAAGTAAATCTCTGTAATAGAAAGCGT 4356  
Db 5131 TAAAAAACCACTATCATGTTATATATTTGCGAAGTAAATCTCTGTAATAGAAAGCGT 5190  
QY 4357 TAGGAACAGATTAGCGAATAATGGGAATAGTTTCAAGTCTCATGTTGATGATCTTGA 4416  
Db 5191 TAGGAACAGATTAGCGAATAATGGGAATAGTTTCAAGTCTCATGTTGATGATCTTGA 5250

4417 TTATATCAAAAGTTTCATATGATATTTTACATCATGAAATCAAGACGCTATTGCTAAGAA 4476  
Db 5251 TTATATCAAAAGTTTCATATGATATTTTACATCATGAAATCAAGACGCTATTGCTAAGAA 5310  
QY 4477 TAAACATATATACGACAAAAAGATATTTTGAACATTAATGATTTTGAATGATGATGATG 4536  
Db 5311 TAAACATATATACGACAAAAAGATATTTTGAACATTAATGATTTTGAATGATGATGATG 5370  
QY 4537 TATACACTGATGATGAAAGCCAAAAAGAGATTTGAAGAAATTTTACCTTTTATGATATAGTGA 4596  
Db 5371 TATACACTGATGATGAAAGCCAAAAAGAGATTTGAAGAAATTTTACCTTTTATGATATAGTGA 5430  
QY 4597 TGACTATATATTTGGTAAATACAAAAGATTTTAAATGCTTTTATTCGCTTTAGGAGGCGGA 4656  
Db 5431 TGACTATATATTTGGTAAATACAAAAGATTTTAAATGCTTTTATTCGCTTTAGGAGGCGGA 5490  
QY 4657 GTTTGGAATTTTAAATACGAATGATGTAAGAAATTTTCAACAACTCTAGCGCTT 4716  
Db 5491 GTTTGGAATTTTAAATACGAATGATGTAAGAAATTTTCAACAACTCTAGCGCTT 5550  
QY 4717 TAGATATATGCTTTAGGCGCAATTTATCAGTGTGATATAGACAAAGTTTATGCAAAAGTTCT 4776  
Db 5551 TAGATATATGCTTTAGGCGCAATTTATCAGTGTGATATAGACAAAGTTTATGCAAAAGTTCT 5610  
QY 4777 TGATGCTGAAACGGGGGAAATTAATGACAAACAAAGAAAGAGTTTATTTGCTGAAAT 4836  
Db 5611 TGATGCTGAAACGGGGGAAATTAATGACAAACAAAGAAAGAGTTTATTTGCTGAAAT 5670  
QY 4837 GAGGAATTAATAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 4896  
Db 5671 GAGGAATTAATAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 5730  
QY 4897 GAAGTTGAATTAAGTACAAATAGATTTATGAGAGGAGGATTTTGAATTAATAA 4956  
Db 5731 GAAGTTGAATTAAGTACAAATAGATTTATGAGAGGAGGATTTTGAATTAATAA 5790  
QY 4957 GCGCCCTTCAGCAAGTTCGCGACTTCCTTTT 4992  
Db 5791 G-CGCCCTTCAGCAAGTTCGCGCGGTTTATTTT 5825

RESULT 10  
US-08-418-085-3/c  
; Sequence 3, Application US/08418085  
; Patent No. 5869283  
; GENERAL INFORMATION:  
; APPLICANT: SLIJKHUIS, HERMAN; SELTEN, GERARDUS CORNELIS  
; APPLICANT: MARIA; SMAAL, ERIC BASTIAN  
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF STEROIDS AND  
; TITLE OF INVENTION: GENETICALLY ENGINEERED CELLS USED THEREIN  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN & MUSERLIAN  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/418,085  
; FILING DATE: 06-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/054,185  
; FILING DATE: 26-APR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/002,608





US-09-099-011A-3/c  
; Sequence 3, Application US/09099011A  
; Patent No. 6,171,836  
; GENERAL INFORMATION:  
; APPLICANT: SLIJKHUIS, HERMAN; SELTEN,  
; APPLICANT: GERARDUS CORNELIS MARIA; SMAAL,  
; APPLICANT: ERIC BASTIAAN  
; TITLE OF INVENTION: PROCESSES FOR OXIDATION OF  
; TITLE OF INVENTION: STEROIDS AND GENETICALLY ENGINEERED CELLS  
; TITLE OF INVENTION: USED THEREIN  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN, MUSERLIAN & LUCAS  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MICROSOFT WORD 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/099,011A  
; FILING DATE: 17-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/418,085  
; FILING DATE: 06-APR-1995  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/054,185  
; FILING DATE: 26-APR-1993  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/002,608  
; FILING DATE: 11-JAN-1993  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/474,857  
; FILING DATE: 30-OCT-1990  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/474,798  
; FILING DATE: 16-JULY-1990  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: PCT/NL89/00072  
; FILING DATE: 25-SEPT-1989  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: NL89/200904.6  
; FILING DATE: 06-MAY-1988  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: NL/88/202080.3  
; FILING DATE: 03-SEP-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CHARLES A. MUSERLIAN  
; REGISTRATION NUMBER: 19,683  
; REFERENCE/DOCKET NUMBER: 146,1169-  
; REFERENCE/DOCKET NUMBER: CON-1-DIV-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 661-8000  
; TELEFAX: (212) 661-8002  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7336 BASE PAIRS  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: DOUBLE  
; TOPOLOGY: UNKNOWN  
; FEATURE:  
; OTHER INFORMATION: PLASMID pSHA-1  
; US-09-099-011A-3  
Query Match 29.0%; Score 1493; DB 3; Length 7336;  
Best Local Similarity 98.1%; Pred. No. 0;  
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
1778 TTCTTATCTGATAATAAGGTAACCTATGCGCGGTAGACTGTAAACATTCTCAGCGATA 1837

Db 5287 AGAATATTGGAGACACCGTTCTTATTTCAGCTATTATAACTCGTCTTCTAAGCATCC 5228  
Qy 2918 TTCAATCCTTTTAAATAAGAAATATAGCATCTAATCTTCAAGAACTGCGGTTTGTTGA 2977  
Db 5227 TTCAATCCTTTTAAATAAGAAATATAGCATCTAATCTTCAAGAACTGCGGTTTGTTGA 5168  
Qy 2978 ACTACTCTTTAATAAATAAATTTTCCGTTCCCAATTCACATTGCAATAATAGAAAATC 3037  
Db 5167 ACTACTCTTTAATAAATAAATTTTCCGTTCCCAATTCACATTGCAATAATAGAAAATC 5108  
Qy 3038 CATCTTCATCGGCTTTTCGTCATCATCTGTATGATCAATCAATTCGCTCTTCTGTGTGAT 3097  
Db 5107 CATCTTCATCGGCTTTTCGTCATCATCTGTATGATCAATCAATTCGCTCTTCTGTGTGAT 5048  
Qy 3098 CAAGGTTTAAATTTTATGATATTCTTTTAAACAAACCAACCATAGAGATTAAACCTTTTAC 3157  
Db 5047 CAAGGTTTAAATTTTATGATATTCTTTTAAACAAACCAACCATAGAGATTAAACCTTTTAC 4988  
Qy 3158 GGTGTAAACCTTCCTCCAAATCAGACAAACAGTTTCAAAATCTTTTCTTCATCATCGGTCA 3217  
Db 4987 GGTGTAAACCTTCCTCCAAATCAGACAAACAGTTTCAAAATCTTTTCTTCATCATCGGTCA 4928  
Qy 3218 TAAATATCGTATCCCTTACAGATATTTTGCAGTTTTCGTCATTCGCGATTGTATATCCG 3277  
Db 4927 TAAATATCGTATCCCTTACAGATATTTTGCAGTTTTCGTCATTCGCGATTGTATATCCG 4868  
Qy 3278 ATTTATATTTATTTTCGTAATTTTATTTTATTTTAAACGTCTCA 3318  
Db 4867 ATTTATATTTATTTTCGTCGAATCAATTTGAACCTTTTACA 4827

## RESULT 12

US-08-470-369-21/c  
; Sequence 21, Application US/08470369  
; Patent No. 6238889  
; GENERAL INFORMATION:  
; APPLICANT: Dorsers, Lambertus C. J.  
; APPLICANT: Wagemaker, Gerard  
; APPLICANT: Vos, Yvonne J.  
; APPLICANT: Van Leen, Robert W.  
; TITLE OF INVENTION: MOLECULAR CLONING AND EXPRESSION OF HUMAN  
; TITLE OF INVENTION: IL-3  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Morrison & Foerster  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,369  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/321,480  
; FILING DATE:  
; APPLICATION NUMBER: 07/854,297  
; FILING DATE: 19-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gracey, Nancy J.  
; REGISTRATION NUMBER: 28,216  
; REFERENCE/DOCKET NUMBER: 24615-20002.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-813-5500  
; TELEFAX: 415-494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 7336 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (plasmid)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; IMMEDIATE SOURCE:  
; CLONE: pBHAL  
; US-08-470-369-21  
Query Match 29.0%; Score 1493; DB 3; Length 7336;  
Best Local Similarity 98.1%; Pred. No. 0;  
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Caps 0;  
Qy 1778 TTCTTATCTTGATAATAAGGGTAACCTATTCGCGGATAGACTGTACATCTCTCACCATA 1837  
Db 6367 TACTCTTTTAAATATCCCGACCTGGCAATGCGGGATAGACTGTACATCTCTCACCATA 6308  
Qy 1838 AAATCCCTTTTCAATTTCTTAATCTTAATCTTATTAATCTTATTAATCAATTCGTCAT 1897  
Db 6307 AAATCCCTTTTCAATTTCTTAATCTTATTAATCTTATTAATCAATTCGTCAT 6248  
Qy 1898 AATTAATCCTTTTCTTATTAATCTTATTAATCTTATTAATCTTATTAATCAATTCGTCAT 1957  
Db 6247 AATTAATCCTTTTCTTATTAATCTTATTAATCTTATTAATCTTATTAATCAATTCGTCAT 6188  
Qy 1958 AATGCGCATGACAGGCATGATAATTAATCTTAATCTTATTAATCTTATTAATCAATTCGTCAT 2017  
Db 6187 AATGCGCATGACAGGCATGATAATTAATCTTATTAATCTTATTAATCAATTCGTCAT 6128  
Qy 2018 AACATGATTAACAAATTTATAGAGTCAATCGTTCATAATGGTATGCGTTTGCACATATCA 2077  
Db 6127 AACATGATTAACAAATTTATAGAGTCAATCGTTCATAATGGTATGCGTTTGCACATATCA 6068  
Qy 2078 CTATATATCCGTCGTTCTGTCCTCACTCCGATCCATTCAGAAATCTCTAGGATTT 2137  
Db 6067 CTATATATCCGTCGTTCTGTCCTCACTCCGATCCATTCAGAAATCTCTAGGATTT 6008  
Qy 2138 CCAGAGTTTCTCAGAGTCGGAAGTTGACAGACATTAAGAACTGGCAGATGTCAT 2197  
Db 6007 CCAGAGTTTCTCAGAGTCGGAAGTTGACAGACATTAAGAACTGGCAGATGTCAT 5948  
Qy 2198 AACCTGAAGAGATCTGATTCGTTTCACTGCTTCACTTAAGACCGAGCGCTCGTAT 2257  
Db 5947 AACCTGAAGAGATCTGATTCGTTTCACTGCTTCACTTAAGACCGAGCGCTCGTAT 5888  
Qy 2258 AACAGATGCGATGATGACAGACCAATCAACATGGCACCTGCTTACCTGTACAGTCA 2317  
Db 5887 AACAGATGCGATGATGACAGACCAATCAACATGGCACCTGCTTACCTGTACAGTCA 5828  
Qy 2318 AGGATGTAGAAATGTTGTCGTCCTTTCGACAGCAATATTAAGCAATTTGCTGCAATTT 2377  
Db 5827 AGGATGTAGAAATGTTGTCGTCCTTTCGACAGCAATATTAAGCAATTTGCTGCAATTT 5768  
Qy 2378 CAACACGCTCTCTACGATAAGGCAAAATCGCATCGTGGAACTGTTGGGCTTCTACCG 2437  
Db 5767 CAACACGCTCTCTACGATAAGGCAAAATCGCATCGTGGAACTGTTGGGCTTCTACCG 5708  
Qy 2438 ATTTAGCAGTTTGATACATCTTCTCTAAGTATCCACTGAATCAATAAATCGGCAAAATAG 2497  
Db 5707 ATTTAGCAGTTTGATACATCTTCTCTAAGTATCCACTGAATCAATAAATCGGCAAAATAG 5648  
Qy 2498 AGAAAAATTGACCATGTGTAAGCGGCAATCTGATTCACCTGAGATGATATCTAGTA 2557  
Db 5647 AGAAAAATTGACCATGTGTAAGCGGCAATCTGATTCACCTGAGATGATATCTAGTA 5588  
Qy 2558 GAATCTCTTCGCTATCAAAATTCACCTTCCACTCCAGCGTTGCTCATTCATTCATGGC 2617  
Db 5587 GAATCTCTTCGCTATCAAAATTCACCTTCCACTCCAGCGTTGCTCATTCATTCATGGC 5528  
Qy 2618 TGAACCTCTCTCTCTGTTGATGACACATCATCTCAATATCCGAATAGGGCCCAT 2677

RESULT 13  
US-09-098-877B-3/c  
; Sequence 3, Application US/09098877B  
; Patent No. 6632633  
; GENERAL INFORMATION:  
; APPLICANT: SLIUKHUIS, HERMAN; SELTEN, GERARDUS CORNELIS  
; APPLICANT: MARIA; SMAAL, ERIC BASTIAN  
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF STEROIDS AND  
; TITLE OF INVENTION: GENETICALLY ENGINEERED CELLS USED THEREIN  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN & MUSERLIAN  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/098.877B  
; FILING DATE:

1778	QY	TTCCTTATCTTGATAAATAAGGTAACCTATTTCGCGGATAGACTGTAACTATCTCAGCAATA	1833
6367	DB	TACTCTTTTAAATATCCCGCAGTGGCAATGCCGGATAGACTGTAACTATCTCAGCAATA	6308
1838	QY	AAATCCCCCTTTCATTTTCTAAATGTAATCTATTACCTTATTATTAAATCGAATTCGCTCAT	1897
6307	DB	AAATCCCCCTTTCATTTTCTAAATGTAATCTATTACCTTATTATTAAATCGAATTCGCTCAT	6248
1898	QY	AATTAATCCCTTTTCTTATTACGCAAAATGGCCCGATTTTAAGCACACCCCTTTATTCGGTT	1957
6247	DB	AATTAATCCCTTTTCTTATTACGCAAAATGGCCCGATTTTAAGCACACCCCTTTATTCGGTT	6188
1958	QY	AATGCGCCATGACAGCGCATGATAAATTACTAACTAGGAGAAGTTAATAAAATCGTAAACC	2017
6187	DB	AATGCGCCATGACAGCGCATGATAAATTACTAACTAGGAGAAGTTAATAAAATCGTAAACC	6128
2018	QY	AACATGATTAACAAATATTAGAGGTCATCGTTCAAAATGGTATGCGTTTTGACACATCCA	2077
6127	DB	AACATGATTAACAAATATTAGAGGTCATCGTTCAAAATGGTATGCGTTTTGACACATCCA	6068
2078	QY	CTATATATCCGTTGCGTTCTGTCCACTCTGGAATCCCATTCAGAAATTTCTCTAGCGATT	2137
6067	DB	CTATATATCCGTTGCGTTCTGTCCACTCTGGAATCCCATTCAGAAATTTCTCTAGCGATT	6008
2138	QY	CCAGAAGTTTCTCAGAGTCGGAAAGTTGACACAGACATTACGAACTGGCACAGATGGTTCAT	2197

Db 6007 CCAGAAATTTCTCAGAGTCGAAAGTTGACACAGACATTAACAACTGGCACAGATGGTCAT 5948  
Qy 2198 AACCTGAAGGAGATCTGATTTCTTAAGTCTGATTAAGACGAGCGCTCGTCGTAT 2257  
Db 5947 AACCTGAAGGAGATCTGATTTCTTAAGTCTGATTAAGACGAGCGCTCGTCGTAT 5888  
Qy 2258 AACAGATGCGATGATGACAGACCAATCAACATGGCACCTGCCATTTGCTACGTGTACAGTCA 2317  
Db 5887 AACAGATGCGATGATGACAGACCAATCAACATGGCACCTGCCATTTGCTACGTGTACAGTCA 5828  
Qy 2318 AGATGGTAGAATGTTGTCGCTCTGACACAGATTAAGCCATTTGCTGCTGATTT 2377  
Db 5827 AGATGGTAGAATGTTGTCGCTCTGACACAGATTAAGCCATTTGCTGCTGATTT 5768  
Qy 2378 CAAACAGCTCTTCTACGATAAGGACACAAATCGCATCGTGGAACTTTGGGCTTCTACCG 2437  
Db 5767 CAAACAGCTCTTCTACGATAAGGACACAAATCGCATCGTGGAACTTTGGGCTTCTACCG 5708  
Qy 2438 ATTTAGCAGTTTGTATACATTTCTTAAGTATCCACCTGAATCATATAATCGCAAAATAG 2497  
Db 5707 ATTTAGCAGTTTGTATACATTTCTTAAGTATCCACCTGAATCATATAATCGCAAAATAG 5648  
Qy 2498 AGAAAAATGACATGTAAGCGGCAATCTGATTCACCTGAGATGATAATCTAGTA 2557  
Db 5647 AGAAAAATGACATGTAAGCGGCAATCTGATTCACCTGAGATGATAATCTAGTA 5588  
Qy 2558 GAATCTTTCGTATCAAAATTCATTCACCTTCACCTGACCGGTTGTCCATTCATGGC 2617  
Db 5587 GAATCTTTCGTATCAAAATTCATTCACCTTCACCTGACCGGTTGTCCATTCATGGC 5528  
Qy 2618 TGAATCTGCTTCTGTTGATGACATGACACATCTCATATCCGAATAGGCCCAT 2677  
Db 5527 TGAATCTGCTTCTGTTGATGACACATCTCATATCCGAATAGGCCCAT 5468  
Qy 2678 CAGTCTGACGACCAAGAGAGCCATAAACCAATAGCCCTTAACATCATCCCATATTTAT 2737  
Db 5467 CAGTCTGACGACCAAGAGAGCCATAAACCAATAGCCCTTAACATCATCCCATATTTAT 5408  
Qy 2738 CCAATATTCGTTTAAATTTTCAAGAACATCTTCAATCTTCTCTAGTCATATTA 2797  
Db 5407 CCAATATTCGTTTAAATTTTCAAGAACATCTTCAATCTTCTCTAGTCATATTA 5348  
Qy 2798 TTGGTCCATCTACTATTCATTCCTTTTCAGATAATTTAGATTTGTTTCTAAATA 2857  
Db 5347 TTGGTCCATCTACTATTCATTCCTTTTCAGATAATTTAGATTTGTTTCTAAATA 5288  
Qy 2858 AGAATATTTGGAGACCGGTTCTTATTCAGCTATTAATACTGCTTCTTCAAGCATCC 2917  
Db 5287 AGAATATTTGGAGACCGGTTCTTATTCAGCTATTAATACTGCTTCTTCAAGCATCC 5228  
Qy 2918 TTCAATCCTTTTAAATTAACATATATGATCTTCAACAACTGGCCCGTTTGTGA 2977  
Db 5227 TTCAATCCTTTTAAATTAACATATATGATCTTCAACAACTGGCCCGTTTGTGA 5168  
Qy 2978 ACTACTCTTTTAAATAAATTTTCCGTTCCCAATTCACATTCGAATTAAGAAAATC 3037  
Db 5167 ACTACTCTTTTAAATAAATTTTCCGTTCCCAATTCACATTCGAATTAAGAAAATC 5108  
Qy 3038 CATCTTCATCGGCTTTTCTGTCATCTGATGATCAATCAATCGCTTCTTCTGTCAT 3097  
Db 5107 CATCTTCATCGGCTTTTCTGTCATCTGATGATCAATCAATCGCTTCTTCTGTCAT 5048  
Qy 3098 CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCAATAGGAGATTAACCTTTTAC 3157  
Db 5047 CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCAATAGGAGATTAACCTTTTAC 4988  
Qy 3158 GGTGTAACCTTCTCCAAATCAGACAAAGTTTCAAAATCTTTCTTCATCATCGTCA 3217  
Db 4987 GGTGTAACCTTCTCCAAATCAGACAAAGTTTCAAAATCTTTCTTCATCATCGTCA 4928  
Qy 3218 TAAATCCGATCCTTTACAGGATATTTTCAGTTTCGTCATTTGCGGATTTGATATCG 3277

Db 4927 TAAAAATCCGATCCCTTTACAGGATATTTTCAGTTTCGATTCGATTCGATTCG 4868  
Qy 3278 ATTTATATTTATTTTCGATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTT 3318  
Db 4867 ATTTATATTTATTTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 4827  
RESULT 14  
US-08-556-978B-79  
; Sequence 79, Application US/08556978B  
; Patent No. 6268169  
; GENERAL INFORMATION:  
; APPLICANT: FARNESOCK, STEPHEN F.  
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED  
; TITLE OF INVENTION: SPIDER SILK ANALOGS  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
; STREET: 1007 MARKET STREET  
; CITY: WILMINGTON  
; STATE: DELAWARE  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 19898  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.50 INCH  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: MICROSOFT WINDOWS 95  
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/556,978B  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/077,600  
; FILING DATE: JUNE 15, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FLOYD, LINDA AXAMETHY  
; REGISTRATION NUMBER: 33,692  
; REFERENCE/DOCKET NUMBER: CR-9389-A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302-892-8112  
; TELEFAX: 302-773-0164  
; INFORMATION FOR SEQ ID NO: 79:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9144 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
US-08-556-978B-79  
Query Match 29.0%; Score 1493; DB 3; Length 9144;  
Best Local Similarity 98.1%; Pred. No. 0;  
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1778 TTCTTATCTCTGATTAATTAAGGTAACTATTCGCGGATAGACTGTAAACATCTCACGCATA 1837  
Db 6526 TACTCTTTTAAATATCCCGACTGGCAATCGCGGATAGACTGTAAACATCTCACGCATA 6585  
Qy 1838 AAATCCCTTTTCATTTTCTTAATGTAATCTATTACCTTTATTTATTTATTTATTTATTTATTT 1897  
Db 6586 AAATCCCTTTTCATTTTCTTAATGTAATCTATTACCTTTATTTATTTATTTATTTATTTATTT 6645  
Qy 1898 AATTAATCTTTTCTTTATTTACGCAAAATGCGCGATTTAAGCACACCTTTATTCGTT 1957  
Db 6646 AATTAATCTTTTCTTTATTTACGCAAAATGCGCGATTTAAGCACACCTTTATTCGTT 6705  
Qy 1958 AATCGCCATGACAGCCATGATTAATTAATCTAGAGAGTTAATAATAGTAACC 2017  
Db 6706 AATCGCCATGACAGCCATGATTAATTAATCTAGAGAGTTAATAATAGTAACC 6765  
Qy 2018 AACATGATTAACAATTTATAGAGTCACTCGTTCAAAATGGTATCGCTTTTGACATCCA 2077



||||| 5827 AGGATGCTAGAAATGTTGTCGGTCTTGCCACAGATATTACGCCATTTGGCTGCATATT 5768  
Db  
||||| 2378 CAAACAGCTCTTCTACGATTAAGGGCAAAATCGCATCGTGGAAAGTTTGGGCTTCTACCG 2437  
Qy  
||||| 5767 CAAACAGCTCTTCTACGATTAAGGGCAAAATCGCATCGTGGAAAGTTTGGGCTTCTACCG 5708  
Db  
||||| 2438 ATTTAGCAGTTTGATACACATTTCTTAAGTATCCACCTGAATCATAAATCGGCAAAATAG 2497  
Qy  
||||| 5707 ATTTAGCAGTTTGATACACATTTCTTAAGTATCCACCTGAATCATAAATCGGCAAAATAG 5648  
Db  
||||| 2498 AGAAAAATTGACCAATGTTGTAAGCGGCCAAATCTGATTCACCTGAGATGCAATATCTAGTA 2557  
Qy  
||||| 5647 AGAAAAATTGACCAATGTTGTAAGCGGCCAAATCTGATTCACCTGAGATGCAATATCTAGTA 5588  
Db  
||||| 2558 GAATCTCTTCGCTATCAAAATTCATCTCCACTTCCACTCACCGTTGTCCATTCATGGC 2617  
Qy  
||||| 5587 GAATCTCTTCGCTATCAAAATTCATCTCCACTTCCACTCACCGTTGTCCATTCATGGC 5528  
Db  
||||| 2618 TGAATCTCTGCTTCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGGCCCAT 2677  
Qy  
||||| 5527 TGAATCTCTGCTTCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGGCCCAT 5468  
Db  
||||| 2678 CAGTCTGACGACGACGAGAGCCGCAATTAACCAATAGGCTTAACATCATCCCCATATTAT 2737  
Qy  
||||| 5467 CAGTCTGACGACGACGAGAGCCGCAATTAACCAATAGGCTTAACATCATCCCCATATTAT 5408  
Db  
||||| 2738 CCAATATTCTGTTTCTTAATTTTCATGAACAACTTTCATCTTCTCTCTAGTCAATTATA 2797  
Qy  
||||| 5407 CCAATATTCTGTTTCTTAATTTTCATGAACAACTTTCATCTTCTCTCTAGTCAATTATA 5348  
Db  
||||| 2798 TTGTTCCATTCACATTTCTTCCCTTTTCCAGATTAATTTAGATTTGCTTTCTTAATA 2857  
Qy  
||||| 5347 TTGTTCCATTCACATTTCTTCCCTTTTCCAGATTAATTTAGATTTGCTTTCTTAATA 5288  
Db  
||||| 2858 AGAATATTGAGAGACCGTTCTTATTCAGCTATTAATTAATCGTCTTCTTCTTAAAGCATCC 2917  
Qy  
||||| 5287 AGAATATTGAGAGACCGTTCTTATTCAGCTATTAATTAATCGTCTTCTTCTTAAAGCATCC 5228  
Db  
||||| 2918 TTCAATCTCTTTTAATAACAATTTATAGCATCTAATCTTCAACAACTGGCCGTTTGTGA 2977  
Qy  
||||| 5227 TTCAATCTCTTTTAATAACAATTTATAGCATCTAATCTTCAACAACTGGCCGTTTGTGA 5168  
Db  
||||| 2978 ACTACTCTTTTAATAACAATTTTCCGTTCCCATTCGCATTCGATTAATAGAAATC 3037  
Qy  
||||| 5167 ACTACTCTTTTAATAACAATTTTCCGTTCCCATTCGCATTCGATTAATAGAAATC 5108  
Db  
||||| 3038 CATCTTCATCGGCTTTTTCGTCATCATCTGTATGAATCAATTCGCTTCTTCTGTGTCAT 3097  
Qy  
||||| 5107 CATCTTCATCGGCTTTTTCGTCATCATCTGTATGAATCAATTCGCTTCTTCTGTGTCAT 5048  
Db  
||||| 3098 CAAGGTTAAATTTTATGATTTCTTTTAACAAACCCATAGGAGATTACCTTTTAC 3157  
Qy  
||||| 5047 CAAGGTTAAATTTTATGATTTCTTTTAACAAACCCATAGGAGATTACCTTTTAC 4988  
Db  
||||| 3158 GGTGTAACCTTCTCCCAATCAGACAAACGTTTCAAAATTCCTTTCTTCATCATCGGTCA 3217  
Qy  
||||| 4987 GGTGTAACCTTCTCCCAATCAGACAAACGTTTCAAAATTCCTTTCTTCATCATCGGTCA 4928  
Db  
||||| 3218 TAAATCCGATCTCTTTACGATATTTTTCAGTTTCGTCATTTGCGAATTCGATTTATCCG 3277  
Qy  
||||| 4927 TAAATCCGATCTCTTTACGATATTTTTCAGTTTCGTCATTTGCGAATTCGATTTATCCG 4868  
Db  
||||| 3278 ATTTATATTATTTTTCGTTATTTTATTAACGCTCTCA 3318  
Qy  
||||| 4867 ATTTATATTATTTTTCGTTGGAATCATTTGAACCTTTTACA 4827  
Db

This Page Blank (usp10)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 02:36:16 ; Search time 1169.56 Seconds  
(without alignments)  
18677.278 Million cell updates/sec

Title: US-10-030-390-1

Perfect score: 5142

Sequence: 1 gaattcgagctcggtaccg.....ccgtctgaagttctcttaag 5142

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 29Jan04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002s: \*  
7: Geneseqn2003as: \*  
8: Geneseqn2003bs: \*  
9: Geneseqn2003cs: \*  
10: Geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5142	100.0	5142	5	Aaf29645 Plasmid p
2	4232.8	82.3	10929	6	Abk98591 pep25 vec
3	4232.8	82.3	10929	8	AcD13842 Plasmid p
C 4	2467.2	48.0	8119	2	Aat33477 Subtilisi
C 5	1650.4	32.1	2059	1	Aan70271 DNA fragm
6	1636.8	31.8	3792	2	AaQ48463 Plasmid p
7	1635.2	31.8	5234	2	AaQ48464 Plasmid p
8	1635.2	31.8	6722	2	AaQ48465 Plasmid p
C 9	1493	29.0	7336	1	Aan91856 Plasmid p
10	1493	29.0	8808	2	Aat32227 Plasmid p
11	1493	29.0	9144	2	Aat18951 Plasmid p
12	1493	29.0	10140	2	Aat32332 Plasmid p
13	1490.6	28.9	6661	7	AaQ49392 Plasmid p
14	1487.6	28.9	5302	3	AaA98012 Expressio
15	1487.6	28.9	5767	3	AaA98014 Expressio
16	1481	28.8	2763	2	AaQ48044 Lactobaci
C 17	1479	28.8	5609	6	AaD23867 pCF717 D
C 18	1479	28.8	5609	6	AaD23868 pEB200 PL
C 19	1479	28.8	7026	6	AaD23871 pEB303 PL
C 20	1479	28.8	7344	6	AaD23869 pEB301 ch
C 21	1476.2	28.7	2763	2	AaQ46037 Plasmid p
C 22	1464	28.5	7534	1	Aan80664 Sequence
C 23	1363	26.5	7456	2	AaQ10686 Plasmid p

## ALIGNMENTS

## RESULT 1

## AAF29645

ID AAF29645 standard; DNA; 5142 BP.

XX AC AAF29645;

DT 10-APR-2001 (first entry)

DE Plasmid pL2mTFF1v1.

XX Mouse TPFF1; mTFF; trefoil protein; antiinflammatory; gastrointestinal;  
anti-ulcer; peptide therapy; gastrointestinal disease; acute colitis;  
Crohn's disease; ulcerative colitis; plasmid pL2mTFF1v1;  
recombinant vector; ds.

XX Synthetic.

OS WO200102570-A1.

PN 11-JAN-2001.

PD 05-JUL-2000; 2000WO-EP006343.

XX 05-JUL-1999; 99EP-00870143.

XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

PI Hans WC, Steidler L, Remaut ER;

XX WPI; 2001-138142/14.

DR Recombinant Lactococcus lactis for delivering a trefoil peptide useful  
for treating acute or chronic gastrointestinal inflammatory diseases or  
disorders, e.g. acute or ulcerative colitis, acute flare-ups of Crohn's  
disease.

XX Claim 18; Fig 1b; 4app; English.

XX The present sequence is a recombinant vector which may be used to  
generate a recombinant lactococcus lactis capable of delivering a trefoil  
peptide in vivo. The recombinant microorganism is useful for  
manufacturing an agent for the delivery of a trefoil peptide to the  
gastrointestinal tract, and for treating gastric or intestinal diseases  
or disorders, or lesions caused by gastric or intestinal diseases or  
disorders. The microorganism may also be used for preparing medicament to  
be used for treating gastric and /or gastrointestinal diseases or

CC disorders, acute gastrointestinal inflammatory diseases (e.g., acute  
 CC colitis, acute flare-ups of Crohn's diseases, or ulcerative colitis), and  
 CC chronic and spontaneously recurring diseases of the gastrointestinal  
 CC tract comprising Crohn's disease (enteritis regionalis) and ulcerative  
 CC colitis (colitis ulcerosa). Disease states which can be treated by the  
 CC method or compositions comprising the recombinant microorganism or  
 CC trefoil peptides include disorders of and damage to the alimentary canal,  
 CC including the mouth, oesophagus, stomach and large and small intestine,  
 CC as well as for the protection and treatment of tissues that lie outside  
 CC the alimentary canal  
 XX

SQ Sequence 5142 BP; 1617 A; 954 C; 927 G; 1644 T; 0 U; 0 Other;

Query Match 100.0%; Score 5142; DB 5; Length 5142;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAATTCGAGCTCGGTACCGGGGATCTCGATCCCGCGAATTAATACGACTCTACTATAGG	60
DB	1	GAATTCGAGCTCGGTACCGGGGATCTCGATCCCGCGAATTAATACGACTCTACTATAGG	60
QY	61	GAGACCAACAGCGTTCCCTCTAGAAATAATTTTGTAACTTTAAGAAGAGATATACA	120
DB	61	GAGACCAACAGCGTTCCCTCTAGAAATAATTTTGTAACTTTAAGAAGAGATATACA	120
QY	121	TATGAAAAAAGATTATCTCAGCTATTTTAATGCTACAGTATCTTCTGCTGACG	180
DB	121	TATGAAAAAAGATTATCTCAGCTATTTTAATGCTACAGTATCTTCTGCTGACG	180
QY	181	CCCGTTGTGAGGTGTTACGCCCGCAGGCCAGGCCCGCAGGAAACATGTATCAT	240
DB	181	CCCGTTGTGAGGTGTTACGCCCGCAGGCCAGGCCCGCAGGAAACATGTATCAT	240
QY	241	GGCCCCCGGAGAGGATTAATTTGGTTCCTCCGGTGTACCGCCCGCAGGTGCGAGGA	300
DB	241	GGCCCCCGGAGAGGATTAATTTGGTTCCTCCGGTGTACCGCCCGCAGGTGCGAGGA	300
QY	301	GAGAGGTGCTGTTTGTATGACAGTGTCCGGGATTCCTGGTGTCTTCCACCCCATGGC	360
DB	301	GAGAGGTGCTGTTTGTATGACAGTGTCCGGGATTCCTGGTGTCTTCCACCCCATGGC	360
QY	361	CATCGAGAACACTCAAGAAGAAATGTCCCTCTAACTAGTAGATCCGCTGCTAACAA	420
DB	361	CATCGAGAACACTCAAGAAGAAATGTCCCTCTAACTAGTAGATCCGCTGCTAACAA	420
QY	421	AGCCGGAAGAGAGCTGAGTGGCTGTCCACCGCTGACCAATTAAGTACATACCCCT	480
DB	421	AGCCGGAAGAGAGCTGAGTGGCTGTCCACCGCTGACCAATTAAGTACATACCCCT	480
QY	481	TGGGGCCCTCTAAACGGGTCTTGGAGGGTTTTTGTCTGAAAGAGGAACTATATCCGGATG	540
DB	481	TGGGGCCCTCTAAACGGGTCTTGGAGGGTTTTTGTCTGAAAGAGGAACTATATCCGGATG	540
QY	541	ACCTGCGAGGATGAAGCTGGGCTGTCCCGCTGTCTTACAGTGTGACTGGGAAAC	600
DB	541	ACCTGCGAGGATGAAGCTGGGCTGTCCCGCTGTCTTACAGTGTGACTGGGAAAC	600
QY	601	CCTGGCGTTACCCAACTTAATTCGGCTTGCAGCACATCCCGCTTTCGCCAGCTCATTTAC	660
DB	601	CCTGGCGTTACCCAACTTAATTCGGCTTGCAGCACATCCCGCTTTCGCCAGCTCATTTAC	660
QY	661	TTTTTGATTTCTACAACTGATCAATGATTAATTCGCTCTTTTATAGTGGGACAA	720
DB	661	TTTTTGATTTCTACAACTGATCAATGATTAATTCGCTCTTTTATAGTGGGACAA	720
QY	721	ATGAGGCGATTTTCGCTCTTTTCGGCGAGCTAGTTACCTTAAGTTATGATGACT	780
DB	721	ATGAGGCGATTTTCGCTCTTTTCGGCGAGCTAGTTACCTTAAGTTATGATGACT	780
QY	781	GGTTTTAAGCGCAAAAAGTTGCTTTTTCGTAACCTATTAATGATCGTTTTAGAAAACC	840
DB	781	GGTTTTAAGCGCAAAAAGTTGCTTTTTCGTAACCTATTAATGATCGTTTTAGAAAACC	840

QY	841	GACTGTAAAAAGTACAGTCGGCATTATCTCATATTATAAAGCCAGTCATTAGGCTATC	900
DB	841	GACTGTAAAAAGTACAGTCGGCATTATCTCATATTATAAAGCCAGTCATTAGGCTATC	900
QY	901	TGACAAATTCCTGAATAGAGTTCATAAACAATCCTGCATGATAACCATCACAAACGAATG	960
DB	901	TGACAAATTCCTGAATAGAGTTCATAAACAATCCTGCATGATAACCATCACAAACGAATG	960
QY	961	ATGTACTCTCTAAAGATAGCGGTAAATATATGAAATACCTTTATTAAATGAATTTCTTCG	1020
DB	961	ATGTACTCTCTAAAGATAGCGGTAAATATATGAAATACCTTTATTAAATGAATTTCTTCG	1020
QY	1021	TGTAATAATGGGTAGAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1080
DB	1021	TGTAATAATGGGTAGAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1080
QY	1081	TGAAGTCCCATGGAATAAATAGAAAGAGAAAAGCAATTTTCAGGTATAGGTGTTTTGGGAAA	1140
DB	1081	TGAAGTCCCATGGAATAAATAGAAAGAGAAAAGCAATTTTCAGGTATAGGTGTTTTGGGAAA	1140
QY	1141	CAATTTCCCGGAACCATTAATTTCTCTACATCAGAAAGGTAAATATCAATAAACTCTTT	1200
DB	1141	CAATTTCCCGGAACCATTAATTTCTCTACATCAGAAAGGTAAATATCAATAAACTCTTT	1200
QY	1201	GAAGTCATTTCTACAGGAGTCCCAATACAGAGAGTATTTTAGATACACCATCAAAAAT	1260
DB	1201	GAAGTCATTTCTACAGGAGTCCCAATACAGAGAGTATTTTAGATACACCATCAAAAAT	1260
QY	1261	TGTAATAAGTGGCTCTAACTTATCCCAATAACCTCTCCGTCGCTATTTGTAACCACT	1320
DB	1261	TGTAATAAGTGGCTCTAACTTATCCCAATAACCTCTCCGTCGCTATTTGTAACCACT	1320
QY	1321	TCATAAAGCTGTATTTGAGTTTATCACCCCTTGTCTCACTAAAGAAAATAATGAGGGTAAA	1380
DB	1321	TCATAAAGCTGTATTTGAGTTTATCACCCCTTGTCTCACTAAAGAAAATAATGAGGGTAAA	1380
QY	1381	TTTATATCTTCTCTGTTTATGTTTCGGTATAAACAACATAATCAATTTCTGTTGTTAT	1440
DB	1381	TTTATATCTTCTCTGTTTATGTTTCGGTATAAACAACATAATCAATTTCTGTTGTTAT	1440
QY	1441	ACTAAAGCTGTATTTGAGTTTCAATAATTAATTAATTAATTAATTAATTAATTAATTAAT	1500
DB	1441	ACTAAAGCTGTATTTGAGTTTCAATAATTAATTAATTAATTAATTAATTAATTAATTAAT	1500
QY	1501	TAAATCAATTTTATTAAGTTTCAATTCGATGCTCTTAATTTTATCTAAAGTGAATTT	1560
DB	1501	TAAATCAATTTTATTAAGTTTCAATTCGATGCTCTTAATTTTATCTAAAGTGAATTT	1560
QY	1561	TAGGAGGCTTACTTGTCTGCTTCTTCTTCAATTAAGATCAATCCCTTTTAAAGTCAATTT	1620
DB	1561	TAGGAGGCTTACTTGTCTGCTTCTTCTTCAATTAAGATCAATCCCTTTTAAAGTCAATTT	1620
QY	1621	ACTGTAAACATTAATATATTTTAAATAATATCCCACTTATCCAAATTTGTTGTTGAA	1680
DB	1621	ACTGTAAACATTAATATATTTTAAATAATATCCCACTTATCCAAATTTGTTGTTGAA	1680
QY	1681	CTAATGGGTGCTTTAGTTTGAAGATTAAGACCAATTTTAAAGTGAATTTTCTGTTT	1740
DB	1681	CTAATGGGTGCTTTAGTTTGAAGATTAAGACCAATTTTAAAGTGAATTTTCTGTTT	1740
QY	1741	TTTTTAAGGATTTGAGGTAGCGAAAATCTTTTCTTCTTATCTTGAATAAAGGGTA	1800
DB	1741	TTTTTAAGGATTTGAGGTAGCGAAAATCTTTTCTTCTTATCTTGAATAAAGGGTA	1800
QY	1801	ACTATGGCGGATAGACTGTAACTTCTACAGCATAAATCCCTTTTCAATTTCTAATG	1860
DB	1801	ACTATGGCGGATAGACTGTAACTTCTACAGCATAAATCCCTTTTCAATTTCTAATG	1860
QY	1861	TAAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1920
DB	1861	TAAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1920
QY	1921	CAAAATGCGCGATTTTAAAGCACACCCCTTTTATTCGTTTATGCGCCTATGACAGCATGATA	1980

Db	1921		CAAAATGGCCGGAATTTAAAGCACACCCCTTTATTCCGTTAATGGCCATGACAGCCATGATA	1980
Qy	1981		ATTACTAAATACCTAGGAGAGAGTTAAATAAATACGTTAAACCAACATGATTAAACAAATTTATTAGAG	2040
Db	1981		ATTACTAAATACCTAGGAGAGAGTTAAATAAATACGTTAAACCAACATGATTAAACAAATTTATTAGAG	2040
Qy	2041		GTCATCGTTCAAAATGGTATGCGTTTTGTGACACATCCATATATATCCGTTGTCGTTCTGTC	2100
Db	2041		GTCATCGTTCAAAATGGTATGCGTTTTGTGACACATCCATATATATCCGTTGTCGTTCTGTC	2100
Qy	2101		CACCTCTGAATCCGATTCAGAAATTTCTCTAGCGATTCCAGAGTTTCTCAGAGTCGGA	2160
Db	2101		CACCTCTGAATCCGATTCAGAAATTTCTCTAGCGATTCCAGAGTTTCTCAGAGTCGGA	2160
Qy	2161		AGTTGACACAGACATTACGAATCGCACAGATGTCATAACCTGAAGGAAGATCTGATTGC	2220
Db	2161		AGTTGACACAGACATTACGAATCGCACAGATGTCATAACCTGAAGGAAGATCTGATTGC	2220
Qy	2221		TTTAATCTGCTTCAAGAGCGCTCGTCTGATTAACAGATGCGATGCAGACCA	2280
Db	2221		TTTAATCTGCTTCAAGAGCGCTCGTCTGATTAACAGATGCGATGCAGACCA	2280
Qy	2281		ATCAACATGCGACCTGCGCATTTGCTGTACAGTCAAGATCGTGAAGAAATGTTGCGT	2340
Db	2281		ATCAACATGCGACCTGCGCATTTGCTGTACAGTCAAGATCGTGAAGAAATGTTGCGT	2340
Qy	2341		CGTTGACACAGAAATTTACGCAATTTGCTGTGCAATTTCAAAACAGCTCTTCTACGATTAAGG	2400
Db	2341		CGTTGACACAGAAATTTACGCAATTTGCTGTGCAATTTCAAAACAGCTCTTCTACGATTAAGG	2400
Qy	2401		GCACAAATCGCATCGTGGACGTTTGGGCTTCTACCGATTTAGCAGTTTGATACATTTTC	2460
Db	2401		GCACAAATCGCATCGTGGACGTTTGGGCTTCTACCGATTTAGCAGTTTGATACATTTTC	2460
Qy	2461		TCTAAGTATCCACCTGTAATCATAAATCGGCAAAATAGAGAAAATTGACCAATGTGAAGC	2520
Db	2461		TCTAAGTATCCACCTGTAATCATAAATCGGCAAAATAGAGAAAATTGACCAATGTGAAGC	2520
Qy	2521		GGCCAAATCTGATTCACCTGAGATGCATAATCTAGTAGAATCTCTTCGCTATCAAAATTC	2580
Db	2521		GGCCAAATCTGATTCACCTGAGATGCATAATCTAGTAGAATCTCTTCGCTATCAAAATTC	2580
Qy	2581		ACTTCCACCTTCCACTCACCGGTTGCCATTCATGGCTGAACTCTGTTCTCTCTGTGAC	2640
Db	2581		ACTTCCACCTTCCACTCACCGGTTGCCATTCATGGCTGAACTCTGTTCTCTCTGTGAC	2640
Qy	2641		ATGACACACATCTCAATTCGCAATAGGGCCCACTCAGTCTGACACCAAGAGACCA	2700
Db	2641		ATGACACACATCTCAATTCGCAATAGGGCCCACTCAGTCTGACACCAAGAGACCA	2700
Qy	2701		TAAACACCAATAGCCTTAAACATCATCCCATATTTATTCCAATTTCCGTTCTTAAATTTCA	2760
Db	2701		TAAACACCAATAGCCTTAAACATCATCCCATATTTATTCCAATTTCCGTTCTTAAATTTCA	2760
Qy	2761		TGAACAAATCTTCAATCTTCTCTCTAGTCAATTAATTTGGTCAATTCATTTCTCAAT	2820
Db	2761		TGAACAAATCTTCAATCTTCTCTCTAGTCAATTAATTTGGTCAATTCATTTCTCAAT	2820
Qy	2821		CCCTTTTCAGATATTTTAGATTTGCTTTTCTTAATAGATATTTTCGAGAGACCGGTC	2880
Db	2821		CCCTTTTCAGATATTTTAGATTTGCTTTTCTTAATAGATATTTTCGAGAGACCGGTC	2880
Qy	2881		TTTATTCAGCTATTAATAACTCGTCTTCTTAAGCATCTTCAATCTTTTAAATTAACAATTA	2940
Db	2881		TTTATTCAGCTATTAATAACTCGTCTTCTTAAGCATCTTCAATCTTTTAAATTAACAATTA	2940
Qy	2941		TAGCATCTTAATCTTCAACAACTGGCCGGTTTGTGAACTCTTTAAATAAAATATTT	3000
Db	2941		TAGCATCTTAATCTTCAACAACTGGCCGGTTTGTGAACTCTTTAAATAAAATATTT	3000
Qy	3001		TTCCGTTCCCAATTCACATTTGCAATAATAGAAAATCCATCTTCATCGGCTTTTTCGTC	3060

3001	TTCCGTTCCCAATTCACATTCGAAATAATAGAAAAATCCATCTTCATCGGCTTTTTTTCGCTCA	3061	TCATCTGTATGAATCAAAATCGCCTTCCTCTGTGTGTCATCAAGGTTTAAATTTTTTATGATATT
3061	TCATCTGTATGAATCAAAATCGCCTTCCTCTGTGTGTCATCAAGGTTTAAATTTTTTATGATATT	3121	TCCTTTTAACAAAACACCATAGGAGATTAACTTTTACGGTGTAAACCTTCCTCCAAATCA
3121	TCCTTTTAACAAAACACCATAGGAGATTAACTTTTACGGTGTAAACCTTCCTCCAAATCA	3181	GACAAACGTTTCAAATTCCTTTCTTCATCATPCGTCATAAAATCCGTAATCCCTTTACAGGA
3181	GACAAACGTTTCAAATTCCTTTCTTCATCATPCGTCATAAAATCCGTAATCCCTTTACAGGA		
3241	TATTTTTCGAGTTTCGTCAAATTCGCCGATGTATATCCGATTATATATTTTTTTCGGTATT	3301	TTTTTATAAACGTCCTCAAATTCGTTCTCTGGAGCGTTTTAGCGTTATATTTTCGTTAGTTA
3301	TTTTTATAAACGTCCTCAAATTCGTTCTCTGGAGCGTTTTAGCGTTATATTTTCGTTAGTTA		
3361	TCGGCATATTCGTTTAAACAGGCGTTATCTGTAGTCGTAAAGCCCTTCGAGCTAGCGTGCT	3421	TTGAGCGGAAGATGTTGTCTGTAGATTATGAAAGCCGATCACTCAATGAAATAATAAGC
3421	TTGAGCGGAAGATGTTGTCTGTAGATTATGAAAGCCGATCACTCAATGAAATAATAAGC		
3481	GCAGCGTCCTTCTATTTTCGGTTTGAGAGAGGCTCAAGGGAGTTTGAGGGAATGAAATTCCTC	3541	TCATGGGTTTCGATTTTAAABATTCGCTCCAAATTTTGCAGCGGTAGCGCTGGAAAAATTT
3541	TCATGGGTTTCGATTTTAAABATTCGCTCCAAATTTTGCAGCGGTAGCGCTGGAAAAATTT		
3601	TTGAAAAAATTTTGGAAATTTGGAAAAAATGGGGGAAAGGAAGCGAAATTTTTCGTTCCGT	3661	TTGAAAAAATTTTGGAAATTTGGAAAAAATGGGGGAAAGGAAGCGAAATTTTTCGTTCCGT
3661	TTGAAAAAATTTTGGAAATTTGGAAAAAATGGGGGAAAGGAAGCGAAATTTTTCGTTCCGT		
3721	CTGGCTCAAGGTTTGAGGGGTTTTTCAATCGCCAAAGAAATCGCCAAAGCTTTTCGCCAAC	3781	GTTTTTTTATAAATCTATATTTAAGTAGCTTTATTCGTTCTTTTATGATTACAAAGTGATA
3781	GTTTTTTTATAAATCTATATTTAAGTAGCTTTATTCGTTCTTTTATGATTACAAAGTGATA		
3841	CACATAATTTTATAAATTTATTGATTGGAGTTTTTTTAAATGGTGATTTTCAAGATCGAAAA	3901	CACATAATTTTATAAATTTATTGATTGGAGTTTTTTTAAATGGTGATTTTCAAGATCGAAAA
3901	AAAGAGTTATGATTTCTCTGCAAAAGAGCAAGATAAAAATTAACAGATATCGCGGAAC	3961	AAAAGAGTTTTCRAAATCTCGGGTTGCGGCTTAGCTATAGAAGAAATGCAAGAAAGG
3961	AAAAGAGTTTTCRAAATCTCGGGTTGCGGCTTAGCTATAGAAGAAATGCAAGAAAGG		
4021	AATCAGAACAAAAAATAAGCGAAAGCTCGCGTTTTTTAGAAGGATACAGATTTTCGCTTA	4081	CTTGTTTTTGTAAAGTAAATATATCATGGCTAATAAATACTAAAGCTTAGAAAAATTCGG
4081	CTTGTTTTTGTAAAGTAAATATATCATGGCTAATAAATACTAAAGCTTAGAAAAATTCGG		

QY	4141	ATTTTATTATATCCTTGAC	CTCAATTCCTAATGAT	TGGAAAGAAAAAT	TGAGAGTTTGGG	4200
Db	4141	ATTTTATTATATCCTTGAC	CTCAATTCCTAATGAT	TGGAAAGAAAAAT	TGAGAGTTTGGG	4200
QY	4201	CGTATCTATGGCTGT	CAGTCCCTTTACACGAT	ATGCGAAAAAAGATAAAGATACATG		4260
Db	4201	CGTATCTATGGCTGT	CAGTCCCTTTACACGAT	ATGCGAAAAAAGATAAAGATACATG		4260
QY	4261	GAATAGTAGTGATGTTAT	ACGAAATGGAAAGAC	CTATPAAAAACCACTATCACGTTAT		4320
Db	4261	GAATAGTAGTGATGTTAT	ACGAAATGGAAAGAC	CTATPAAAAACCACTATCACGTTAT		4320
QY	4321	ATATATTGCACGAAAT	CCTCTTAACAATAGAAAG	CGTTAGGAAACAAGATTAAAGCGGAAAAAT		4380
Db	4321	ATATATTGCACGAAAT	CCTCTTAACAATAGAAAG	CGTTAGGAAACAAGATTAAAGCGGAAAAAT		4380
QY	4381	GGGGAATAGTTCAGTT	GCCTCATGTTGAGATAC	TTGATTATCAAGAGTTTCATATGNAAT		4440
Db	4381	GGGGAATAGTTCAGTT	GCCTCATGTTGAGATAC	TTGATTATCAAGAGTTTCATATGNAAT		4440
QY	4441	TTTGACTCATGAATCAA	AGGACGCTATGCTAAG	AATAAACAATATACGACAAAAAAGA		4500
Db	4441	TTTGACTCATGAATCAA	AGGACGCTATGCTAAG	AATAAACAATATACGACAAAAAAGA		4500
QY	4501	TATTTTGAACATTAA	TGATTTGATTTGAC	CCCTATATAACACTTGTATCAAGAGCCAAA		4560
Db	4501	TATTTTGAACATTAA	TGATTTGATTTGAC	CCCTATATAACACTTGTATCAAGAGCCAAA		4560
QY	4561	AAGAGAATTGAAGA	ATTTACTTTTAGATATAG	TGGATGACTATAAATTTGTTAAATACAAA		4620
Db	4561	AAGAGAATTGAAGA	ATTTACTTTTAGATATAG	TGGATGACTATAAATTTGTTAAATACAAA		4620
QY	4621	AGATTTAATGCGCTTT	TATTCGCCTTAGGAG	CGGAGTTTGTGAATTTTAAATACGAATGA		4680
Db	4621	AGATTTAATGCGCTTT	TATTCGCCTTAGGAG	CGGAGTTTGTGAATTTTAAATACGAATGA		4680
QY	4681	TGTAAGAATATGTTT	CAACAACTCTAGCG	CTTTAGATTATGTTTGAAGGCAATTA		4740
Db	4681	TGTAAGAATATGTTT	CAACAACTCTAGCG	CTTTAGATTATGTTTGAAGGCAATTA		4740
QY	4741	TCAGTGTGGATATAG	ACAAAGTTATCAAAAG	TTCTGTGCTGAAACGGGGGAAATAAA		4800
Db	4741	TCAGTGTGGATATAG	ACAAAGTTATCAAAAG	TTCTGTGCTGAAACGGGGGAAATAAA		4800
QY	4801	ATGACAAACAAAGAA	AGATTATTTGCTGAAA	ATGAGGAATPAAAAAAGAAATTAAG		4860
Db	4801	ATGACAAACAAAGAA	AGATTATTTGCTGAAA	ATGAGGAATPAAAAAAGAAATTAAG		4860
QY	4861	GACTTAAAGAGCGTAT	TGAAAAGATACAGAGA	ATGGAAGTTGAAATTAAGTACAAACAATA		4920
Db	4861	GACTTAAAGAGCGTAT	TGAAAAGATACAGAGA	ATGGAAGTTGAAATTAAGTACAAACAATA		4920
QY	4921	GATTTATTGAGAGG	AGGATTTATGAATAAT	TAARAGCCCCCTGACGAAGTCGCGACT		4980
Db	4921	GATTTATTGAGAGG	AGGATTTATGAATAAT	TAARAGCCCCCTGACGAAGTCGCGACT		4980
QY	4981	TCGTTCTTTTTTAC	CTCTCGGTTATGAGT	TAGTTCAAATTCGTTCTTTTTTAGGTTCTAA		5040
Db	4981	TCGTTCTTTTTTAC	CTCTCGGTTATGAGT	TAGTTCAAATTCGTTCTTTTTTAGGTTCTAA		5040
QY	5041	ATCGTGTTTTTTCT	TGGAATTCGCTGTTT	TATCCTTTACCTTCTCTACAAACCCCTTAA		5100
Db	5041	ATCGTGTTTTTTCT	TGGAATTCGCTGTTT	TATCCTTTACCTTCTCTACAAACCCCTTAA		5100
QY	5101	AACGTTTTTAAAG	CTTTTTAAGCGCT	GTGACGTTCTCTTAAG	5142	
Db	5101	AACGTTTTTAAAG	CTTTTTAAGCGCT	GTGACGTTCTCTTAAG	5142	

XX	ABK98591;
AC	
CC	07-AUG-2003 (revised)
DT	21-OCT-2002 (first entry)
DT	
DE	PEP25 vector fragment containing XylR/Xylo/CP25 sequences.
DE	
XX	ds; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;
KW	P59; P1P2; PL; xylO; tetO; trpO; malO; lambdaclO; cellular proliferation;
KW	antibiotic; vector.
KW	
XX	
OS	Eubacteria
OS	Bacteriophage lambda.
OS	Escherichia coli.
OS	Synthetic.
XX	
XX	WO200251982-A2.
PN	
PD	04-JUL-2002.
XX	
XX	21-DEC-2001; 2001WO-US050250.
PF	
XX	
XX	27-DEC-2000; 2000US-0259434P.
PR	
XX	06-SEP-2001; 2001US-0048993.
PR	
XX	
PA	(ELIT-) ELITRA PHARM INC.
XX	
XX	Haselbeck R, Wall D, Gross M;
PI	
XX	WPI; 2002-575374/61.
DR	
XX	
XX	Isolated nucleic acid comprises bacterial promoters modified to have
PT	altered activity in at least one gram-positive organism, e.g. Bacillus
PT	anthracis or Clostridium botulinum, useful for regulating gene expression
PT	in bacteria.
XX	
PS	Example 1; Page 202-205; 246pp; English.
XX	
XX	The invention relates to an isolated nucleic acid comprising a fusion
CC	promoter comprising at least one promoter that is modified to have
CC	altered activity in at least one gram-positive organism, or comprising
CC	T5, CP25, P32, P1P2 or PL linked to at least one operator consisting
CC	of xylO, tetO, trpO, malO or lambdaclO, where at least one operator is
CC	positioned so binding of a repressor to an operator represses
CC	transcription from the fusion promoter. Also included are vectors and
CC	host cells comprising the fusion promoters, a method of identifying genes
CC	involved in cellular proliferation or required for proliferation of a
CC	prokaryotic cell using the vector, a method of identifying compounds that
CC	inhibit the proliferation of a prokaryotic cell using the vector, a
CC	method of identifying a compound that reduces the activity or level of a
CC	gene product required for proliferation of a cell using the vector, a
CC	compound identified by the methods, a method of inhibiting the activity
CC	or expression of a gene in an operon required for proliferation using the
CC	vector, manufacturing an antibiotic comprising using the vector or cell
CC	and identifying a nucleic acid with promoter activity in Enterococcus
CC	faecalis. The fusion promoters are useful for regulating nucleic acid or
CC	polypeptide expression, particularly for regulating gene expression in
CC	bacteria and for identifying proliferation-required genes or molecules
CC	with potential antibiotic activity. The modified promoters are also
CC	useful for replacing endogenous promoters to create cells with specific
CC	regulatable genes. The present sequence is vector (or fragment)
CC	incorporating a fusion promoter sequence of the invention. (Updated on 07
CC	-AUG-2003 to correct OS field.)
CC	



QY 2870 GAGCACCCTTCTTATTGAGCTATTAAATAACCTGCTTCCCTTAAGCATCCTTCAATCCTTTT 2929  
DB |||||  
QY 8709 GAGCACCCTTCTTATTGAGCTATTAAATAACCTGCTTCCCTTAAGCATCCTTCAATCCTTTT 8768  
DB |||||  
QY 2930 AATAACAATATTAGCATCTAATCTTCAACAACTGGCCCGTTTCTGTAAGTACTCTTTAA 2989  
DB |||||  
QY 8769 AATAACAATATTAGCATCTAATCTTCAACAACTGGCCCGTTTCTGTAAGTACTCTTTAA 8828  
DB |||||  
QY 2990 TAAAAATATTTTCGGTCCCAATTCACATGCAATGCAATAGAAAAATCCATCTTCATCGG 3049  
DB |||||  
QY 8829 TAAAAATATTTTCGGTCCCAATTCACATGCAATGCAATAGAAAAATCCATCTTCATCGG 8888  
DB |||||  
QY 3050 CTTTTTCGTCACTCTGTATGAATCAAAATCGCCTTCTCTGTCTCATCAAGGTTTAAT 3109  
DB |||||  
QY 8889 CTTTTTCGTCACTCTGTATGAATCAAAATCGCCTTCTCTGTCTCATCAAGGTTTAAT 8948  
DB |||||  
QY 3110 TTTTATGATTTCTTTTAAACAAACCAATAGGAGATTAACTTTTACGGTGTAAACCTT 3169  
DB |||||  
QY 8949 TTTTATGATTTCTTTTAAACAAACCAATAGGAGATTAACTTTTACGGTGTAAACCTT 9008  
DB |||||  
QY 3170 CTTCCAAATCAGACAAACGTTTCAAATCTTTTCTTCTCATCATCGTGCATATAAATCCGTAT 3229  
DB |||||  
QY 9009 CTTCCAAATCAGACAAACGTTTCAAATCTTTTCTTCTCATCATCGTGCATATAAATCCGTAT 9068  
DB |||||  
QY 3230 CTTTACAGGATTTTTCAGCTTTTCTGATTTGCGGATTTGATATCCGATTTATATTTAT 3289  
DB |||||  
QY 9069 CTTTACAGGATTTTTCAGCTTTTCTGATTTGCGGATTTGATATCCGATTTATATTTAT 9128  
DB |||||  
QY 3290 TTTTCGGT--ATTTTATTTAAACGCTCAAATCGTTTCTGGAGCTTTTAGCGTTTAA 3347  
DB |||||  
QY 9129 TTTTCGGTFCGATTTTATTTAAACGCTCAAATCGTTTCTGGAGCTTTTAGCGTTTAA 9188  
DB |||||  
QY 3348 TTTTCGGTATTGCGGATTAATCGTTTAAACAGCGCTTATCGTAGCGTTAAAGCCCTTG 3407  
DB |||||  
QY 9189 TTTTCGGTATTGCGGATTAATCGTTTAAACAGCGCTTATCGTAGCGTTAAAGCCCTTG 9248  
DB |||||  
QY 3408 AGCGTAGCGT--GCTTTGAGCGAAGATGTTGCTGTTAGATTATGAAGCGGATGACTGA 3466  
DB |||||  
QY 9249 AGCGTAGCGGCTTTGAGCGAAGATGTTGCTGTTAGATTATGAAGCGGATGACTGA 9308  
DB |||||  
QY 3467 ATGAATTAATAAGCGCAGCTCTCTATTTTCGGTTGAGGAGCTCAAGGGAGTTTGAG 3526  
DB |||||  
QY 9309 ATGAATTAATAAGCGCAGCTCTCTATTTTCGGTTGAGGAGCTCAAGGGAGTTTGAG 9368  
DB |||||  
QY 3527 GGAATGAATTCCTCATGGTTTGATTTTAAAAATTCGTTGCAATTTTCCGAGCGGTA 3586  
DB |||||  
QY 9369 GGAATGAATTCCTCATGGTTTGATTTTAAAAATTCGTTGCAATTTTCCGAGCGGTA 9428  
DB |||||  
QY 3587 GCGCTGGAAATTTTGAAGAAATTTGGAATTTGGAAGAAAT--GGGGGGAAGGAAGC 3645  
DB |||||  
QY 9429 GCGCTGGAAATTTTGAAGAAATTTGGAATTTGGAAGAAATTTGGGGGGAAGGAAGC 9488  
DB |||||  
QY 3646 GAAATTTTTCCTTCTGCTTACGACCCCTTAAAGTGGCGAGTGCCTAATTTTGTGCCAA 3705  
DB |||||  
QY 9489 GAAATTTTTCCTTCTGCTTACGACCCCTTAAAGTGGCGAGTGCCTAATTTTGTGCCAA 9548  
DB |||||  
QY 3706 AACGCTCTATCCCACTGGCTCAAGGTTTGAAGGTTTTCATCGCCCAACGAATGCC 3765  
DB |||||  
QY 9549 AACGCTCTATCCCACTGGCTCAAGGTTTGAAGGTTTTCATCGCCCAACGAATGCC 9608  
DB |||||  
QY 3766 AACGTTTTCGCGCAACGTTTATATAATCTATATTAAAGTACGTTTATTTGTTTAT 3825  
DB |||||  
QY 9609 AACGTTTTCGCGCAACGTTTATATAATCTATATTAAAGTACGTTTATTTGTTTAT 9668  
DB |||||  
QY 3826 GATTACAAAGTGATACACTAATTTTAAATTTATTTGATTTGAGTTTAAATGCTGA 3885  
DB |||||  
QY 9669 GATTACAAAGTGATACACTAATTTTAAATTTATTTGATTTGAGTTTAAATGCTGA 9728  
DB |||||  
QY 3886 TTTTCAGAAATCGAAAAAAGAGTTATGATTTCTGACAAAGAGCAGATATAAATTTAA 3945  
DB |||||  
QY 9729 TTTTCAGAAATCGAAAAAAGAGTTATGATTTCTGACAAAGAGCAGATATAAATTTAA 9788  
DB |||||  
QY 3946 CAGATATGGCGAAACAAAAAGGTTTTCAAAAATCTGCGG--TTGCGCGGTTAGCTATAGA 4003

DB 9789 CAGATATGGCGAAACAAAAAGGTTTTCAAAAATCTGCGGTTGCGGCTTAGCTATAGA 9848  
QY |||||  
QY 4004 AGAATATGCAAGAAAGGAATCAGAAACAAAAAATAAGCGAAAGCTCCGTTTTTAGAAG 4063  
DB |||||  
QY 9849 AGAATATGCAAGAAAGGAATCAGAAACAAAAAATAAGCGAAAGCTCCGTTTTTAGAAG 9908  
DB |||||  
QY 4064 GATAAGGTTTTTCGCTACTCTGTTTTTGTATAGGTAA--TATATCATGGCTATTAAATATAC 4122  
DB |||||  
QY 9909 GATAAGGTTTTTCGCTACTCTGTTTTTGTATAGGTAAATATATCATGGCTATTAAAAATAC 9968  
DB |||||  
QY 4123 TAAAGCTAGAAATTTTGGATTTTATATATCTCTGACTCAATTCCTAATGATGGAAGA 4182  
DB |||||  
QY 9969 TAAAGCTAGAAATTTTGGATTTTATATATCTCTGACTCAATTCCTAATGATGGAAGA 10028  
DB |||||  
QY 4183 AAAATTAGAGGTTTTGGCGGTATCTATGGCTGTCTGCTTTTACACGATATGGACGAAA 4242  
DB |||||  
QY 10029 AAAATTAGAGGTTTTGGCGGTATCTATGGCTGTCTGCTTTTACACGATATGGACGAAA 10088  
DB |||||  
QY 4243 AAAAGATAAAGATACATGGAATAGTATGATGTTTATACGAAATGGAAGCACTATAAAAA 4302  
DB |||||  
QY 10089 AAAAGATAAAGATACATGGAATAGTATGATGTTTATACGAAATGGAAGCACTATAAAAA 10148  
DB |||||  
QY 4303 ACCACATCTACGTTTATATATTTGACGAAATCTCTGTAACTAATAGAAAGCTTAGGAA 4362  
DB |||||  
QY 10149 ACCACATCTACGTTTATATATTTGACGAAATCTCTGTAACTAATAGAAAGCTTAGGAA 10208  
DB |||||  
QY 4363 CAAGATTAAGCGAAATTTGGGGAATAGTTCAGTTGCTCATGTTGAGATACCTTTATAT 4422  
DB |||||  
QY 10209 CAAGATTAAGCGAAATTTGGGGAATAGTTCAGTTGCTCATGTTGAGATACCTTTATAT 10268  
DB |||||  
QY 4423 CAAGGTTTCATATGAATTTTGAATCAATCAAAAGGACGCTATGCTAAGATAAACA 4482  
DB |||||  
QY 10269 CAAGGTTTCATATGAATTTTGAATCAATCAAAAGGACGCTATGCTAAGATAAACA 10328  
DB |||||  
QY 4483 TATATACGACAAAGAAAGATATTTGAACATTAATGATTTTGTATGACCTATATAAC 4542  
DB |||||  
QY 10329 TATATACGACAAAGAAAGATATTTGAACATTAATGATTTTGTATGACCTATATAAC 10388  
DB |||||  
QY 4543 ACTTGATGAAAGCCAAAGAAAGAGAAATTTGAAGAAATTTTACTTTTAGATATAGTGGATGACTA 4602  
DB |||||  
QY 10389 ACTTGATGAAAGCCAAAGAAAGAGAAATTTGAAGAAATTTTACTTTTAGATATAGTGGATGACTA 10448  
DB |||||  
QY 4603 TAAATTTGTAATACAAAGATTTTAAAGCTTTTATTCGCCCTTAGGGAGCGGAGTTGG 4662  
DB |||||  
QY 10449 TAAATTTGTAATACAAAGATTTTAAAGCTTTTATTCGCCCTTAGGGAGCGGAGTTGG 10508  
DB |||||  
QY 4663 AATTTTAAATACGAATGATGTAAAGAGATTTGTTTCAACAACTCTAGCGCTTTAGATT 4722  
DB |||||  
QY 10509 AATTTTAAATACGAATGATGTAAAGAGATTTGTTTCAACAACTCTAGCGCTTTAGATT 10568  
DB |||||  
QY 4723 ATGTTTGAAGGCAATATCATGTTGATATAGAGCAAGTTATGCAAGGTTCTTGATGC 4782  
DB |||||  
QY 10569 ATGTTTGAAGGCAATATCATGTTGATATAGAGCAAGTTATGCAAGGTTCTTGATGC 10628  
DB |||||  
QY 4783 TGAAGCGGGGAAATTAATAATGACAAACAAAGAAAGAGTTTATGCTGAAAAATGAGGAA 4842  
DB |||||  
QY 10629 TGAAGCGGGGAAATTAATAATGACAAACAAAGAAAGAGTTTATGCTGAAAAATGAGGAA 10688  
DB |||||  
QY 4843 TTAAGAAAGAAATTAAGCACTTTAAAGAGCGTATGAAAGATACAGAGAAATGGAAGTT 4902  
DB |||||  
QY 10689 TTAAGAAAGAAATTAAGCACTTTAAAGAGCGTATGAAAGATACAGAGAAATGGAAGTT 10748  
DB |||||  
QY 4903 GAATTAAGTACAACTAGATTTATGAGAGGAGGATTTATGAAATAAATAAGCCCCC 4962  
DB |||||  
QY 10749 GAATTAAGTACAACTAGATTTATGAGAGGAGGATTTATGAAATAAATAAGCCCCC 10808  
DB |||||  
QY 4963 CTGACGAAGTCCGACCTTCGTTCTTTT 4992  
DB |||||  
QY 10809 CTGACGAAGTCCGAGGGGCTTTTATTTT 10838  
DB |||||







```
Qy 4064 GATACGAGTTTTCCTACTTGTGTTTTCGATAGTAA-TATATCATGCTATTTAAATAATAC 4122
Db 9909 GATACGAGTTTTCCTACTTGTGTTTTCGATAGTAA-TATATCATGCTATTTAAATAATAC 9968
Qy 4123 TAAGCTAGAAATTTTCGATTTTATATATCTGCTCAATCCCTAAATGATTGGAAGA 4182
Db 9969 TAAAGCTAGAAATTTTCGATTTTATATATCTGCTCAATCCCTAAATGATTGGAAGA 10028
Qy 4183 AAAATTAGAGAGTTTGGCGGTATCTATGCTGTCAGTCCCTTTACACGATATGACGAAAA 4242
Db 10029 AAAATTAGAGAGTTTGGCGGTATCTATGCTGTCAGTCCCTTTACACGATATGACGAAAA 10088
Qy 4243 AAAAGATAAGATACATGGAATAGTAGTATGTTATACGAAATGGAAGCAGCTATAAAA 4302
Db 10089 AAAAGATAAGATACATGGAATAGTAGTATGTTATACGAAATGGAAGCAGCTATAAAA 10148
Qy 4303 ACCACATATCACGTTATATATATGACGAAATCCCTGTAACTAATAGAAAGCGCTAGGAA 4362
Db 10149 ACCACATATCACGTTATATATGACGAAATCCCTGTAACTAATAGAAAGCGCTAGGAA 10208
Qy 4363 CAAGATTAAAGCGAAATTTGGGGAATAGTTTCACTGCTCATGTTGAGATACCTGATATAT 4422
Db 10209 CAAGATTAAAGCGAAATTTGGGGAATAGTTTCACTGCTCATGTTGAGATACCTGATATAT 10268
Qy 4423 CAAGAGTTTCAATGAATATTTGACTCATGAATCAAGGACGCTATTCCTAAGAAATAACA 4482
Db 10269 CAAGAGTTTCAATGAATATTTGACTCATGAATCAAGGACGCTATTCCTAAGAAATAACA 10328
Qy 4483 TATATACGACAAAAAAGATATTTTGAACATTAATGATTTGATATGACCGCTATAAAC 4542
Db 10329 TATATACGACAAAAAAGATATTTTGAACATTAATGATTTGATATGACCGCTATAAAC 10388
Qy 4543 ACTTGATGAAGCCAAAGAAAGAGATTCAGAAATTTACTTTTATAGATAGTGGATGACTA 4602
Db 10389 ACTTGATGAAGCCAAAGAAAGAGATTCAGAAATTTACTTTTATAGATAGTGGATGACTA 10448
Qy 4603 TAAATTTGTAATAACAAAGATTTTAAATGCTTTTATTCGCTTGGGAGCGAGTTTGG 4662
Db 10449 TAAATTTGTAATAACAAAGATTTTAAATGCTTTTATTCGCTTGGGAGCGAGTTTGG 10508
Qy 4663 AATTTTAAATACGAATGATTAAGATATTTGTTTCAACAACTCTAGCGCTTTAGATT 4722
Db 10509 AATTTTAAATACGAATGATTAAGATATTTGTTTCAACAACTCTAGCGCTTTAGATT 10568
Qy 4723 ATGTTTTCAGGCGCAATTCAGTGTGGATATAGAGCAAGTTTATGCAAGGTTCTTGTATGC 4782
Db 10569 ATGTTTTCAGGCGCAATTCAGTGTGGATATAGAGCAAGTTTATGCAAGGTTCTTGTATGC 10628
Qy 4783 TGAACCGGGGGAATTAATAATGACAAACAAAGAAAGAGTTTATTCGTGAAATGAGGAA 4842
Db 10629 TGAACCGGGGGAATTAATAATGACAAACAAAGAAAGAGTTTATTCGTGAAATGAGGAA 10688
Qy 4843 TTAATAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATGGAAGTT 4902
Db 10689 TTAATAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATGGAAGTT 10748
Qy 4903 GAATTAAGTACAAATAGATTTATGAGAGAGGAGTTTATGATTAATAATAAGCCCCC 4962
Db 10749 GAATTAAGTACAAATAGATTTATGAGAGAGGAGTTTATGATTAATAATAAGCCCCC 10808
Qy 4963 CTGACGAAGTGCAGACTTCGTTCTTTT 4992
Db 10809 CTGACGAAGTGCAGAGGGGCTTTTATTTT 10838
```

## RESULT 4

AAT33477/c

ID AAT33477 standard; DNA; 8119 BP.

XX

AC AAT33477;

XX

DT 16-DEC-1996 (first entry)

XX

```
DE Subtilisin N62D/G166D double mutant DNA in phagemid pSS5.
XX Subtilisin BPN'; protease; enzyme engineering; protein engineering;
KW Bacillus amyloliquefaciens; mutagenesis; substrate specificity; ss.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT misc_difference 7 /*tag= a
FT /note= "base n at position 7 is not identified in the
FT specification"
FT CDS 103..1251
FT /*tag= b
FT sig_peptide 103..192
FT /*tag= c
FT mat_peptide 424..1248
FT /*tag= d
FT misc_difference 1971 /*tag= e
FT /note= "base n at position 1971 is not identified in the
FT specification"
FT misc_difference 1985 /*tag= f
FT /note= "base n at position 1985 is not identified in the
FT specification"
FT misc_difference 8119 /*tag= g
FT /note= "base n at position 8119 is not identified in the
FT specification"
XX WO9627671-A1.
XX 12-SEP-1996. 96WO-US002861.
XX 27-FEB-1996; 96WO-US002861.
XX 03-MAR-1995; 95US-00398028.
XX 19-JUL-1995; 95US-00504265.
XX (GETH ) GENENTECH INC.
XX Ballinger MD, Wells JA;
XX WPI; 1996-425431/42.
XX P-PSDB; AAW00247.
XX Subtilisin variants for cleaving substrates contg. basic residues - allow
XX effective cleavage of fusion proteins with basic linker sequences.
XX Disclosure; Page 29-35; 83pp; English.
XX A DNA sequence (AAT33476) in phagemid pSS5 codes for the prepro form
XX (AAW00247) of subtilisin N62D/G166D double mutant. The amino acid
XX substitutions alter the substrate specificity of the enzyme, such that it
XX specifically cleaves protein substrates contg. basic amino acid residues
XX at positions P1 and P2. Such an enzyme is useful for cleaving fusion
XX proteins with basic substrate linkers, esp. fusion proteins contg. a
XX dibasic sequence, and for processing hormones or other proteins in vitro
XX or in vivo that contain basic cleavage sites. The DNA sequence can be
XX obtd. by site-directed mutagenesis of the Bacillus amyloliquefaciens
XX subtilisin BPN' gene, and used to produce the variant protease in
XX transformed host cells
XX
SQ Sequence 8119 BP; 2459 A; 1628 C; 1811 G; 2217 T; 0 U; 4 Other;
```

## Query Match

Best Local Similarity 48.0%; Score 2467.2; DB 2; Length 8119;

Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 831 TTGAAAAACCGACTGTAAAAAGTACAGTCGGCGGATTTCTCATATTATAAAGCCATCAT 890

Db 7847 TTGAAAAACCGACTGTAAAAAGTACAGTCGGCGGATTTCTCATATTATAAAGCCATCAT 7788

QY	891	TAGGCTATCTGACAAATTCCTGTAATGAGTTTCATAAAACAATCCTGCAATGATAACCATCAC	950	1971	AGCCATGATAAATTAATAATACTAGGAGAAGTTAATAAATACGTAACCAACATGATTAAACA	2030
DB	7787	TAGGCTATCTGACAAATTCCTGTAATGAGTTTCATAAAACAATCCTGCAATGATAACCATCAC	7728	DB	AGCCATGATAAATTAATAATACTAGGAGAAGTTAATAAATACGTAACCAACATGATTAAACA	6648
QY	951	AAACAGAAATGATGACCTGTAAGATAGCGGTAAATATATGTAATACCTTTTAATGA	1010	QY	ATTATATAGAGTTCATCGTTCAAAATGGTATCGTTTTCACACATCCACTATATATCCGTG	2090
DB	7727	AAACAGAAATGATGACCTGTAAGATAGCGGTAAATATATGTAATACCTTTTAATGA	7668	DB	ATTATATAGAGTTCATCGTTCAAAATGGTATCGTTTTCACACATCCACTATATATCCGTG	6588
QY	1011	ATTTTCCTGCTGTAATATGAGTGAAGGTAAATTAATTAATGATATTAATGATTA	1070	QY	TCGTTCTGTCCACTCCTCAATCCCAATTCAGAAATTCCTTAGCGATTCCAGAAATTTCTC	2150
DB	7667	ATTTTCCTGCTGTAATATGAGTGAAGGTAAATTAATTAATGATATTAATGATTA	7608	DB	TCGTTCTGTCCACTCCTCAATTCAGAAATTCCTTAGCGATTCCAGAAATTTCTC	6528
QY	1071	ACCCAGTAAATGCAAGTCCATCGAATATAGAAAGAGAAAGCAATTTTCAGGTATAGGTG	1130	QY	AGAGTCGGAAGTTGACCCAGACATTAACGAACTGGCAGAGATGGTCAATAACCTGAAGGAAG	6468
DB	7607	ACCCAGTAAATGCAAGTCCATCGAATATAGAAAGAGAAAGCAATTTTCAGGTATAGGTG	7548	DB	AGAGTCGGAAGTTGACCCAGACATTAACGAACTGGCAGAGATGGTCAATAACCTGAAGGAAG	6468
QY	1131	TTTTGGGAAACAATTTCCCGAAGCAATTAATTTCTCTACATCAGAAAGGTATTAATCAT	1190	QY	ATCTGATTTGCTTAACCTGCTTCAAGTTCAGTTTAAGACCGAAGCGCTCGTGTATAACAGATGCCGATG	2270
DB	7547	TTTTGGGAAACAATTTCCCGAAGCAATTAATTTCTCTACATCAGAAAGGTATTAATCAT	7488	DB	ATCTGATTTGCTTAACCTGCTTCAAGTTCAGTTTAAGACCGAAGCGCTCGTGTATAACAGATGCCGATG	6408
QY	1191	AAAACCTCTTTCAAGTCAATCTTTTACAGGAGTCCAAATACAGAGAAATGTTTTAGATACAC	1250	QY	ATGCGAGCAAAATCAACATGGCACCTGCCATTTGCTACCTGTACAGTCAAGGATGGTAGAAA	6348
DB	7487	AAAACCTCTTTCAAGTCAATCTTTTACAGGAGTCCAAATACAGAGAAATGTTTTAGATACAC	7428	DB	ATGCGAGCAAAATCAACATGGCACCTGCCATTTGCTACCTGTACAGTCAAGGATGGTAGAAA	6348
QY	1251	CATCAAAAATTTGATTAAGTGGCTCTAACTTATCCCAATTAACCTPAATCTCCGTCGCTAT	1310	QY	TACGATAAGGGGACAAAATCGCATCGTGGAAAGCTTTGGGCTTCTACCGATTTAGCAGTTTG	2450
DB	7427	CATCAAAAATTTGATTAAGTGGCTCTAACTTATCCCAATTAACCTPAATCTCCGTCGCTAT	7368	DB	TACGATAAGGGGACAAAATCGCATCGTGGAAAGCTTTGGGCTTCTACCGATTTAGCAGTTTG	6228
QY	1311	TGTAAACAGTTTCTAAAGCTGTAATTTGAGTTTATCACCTTGTCACATAAGAAATTAATG	1370	QY	TGTTCTGCGTCTTTCACACGCAATATTTACGCCATTTGCTGCTATATTCACACACGCTCTTC	2390
DB	7367	TGTAAACAGTTTCTAAAGCTGTAATTTGAGTTTATCACCTTGTCACATAAGAAATTAATG	7308	DB	TGTTCTGCGTCTTTCACACGCAATATTTACGCCATTTGCTGCTATATTCACACACGCTCTTC	6288
QY	1371	CAGGTAAATTTATATCTTTCTTTTATGTTTTCGGTATTAACCACTAATATCAATTT	1430	QY	TACGATAAGGGGACAAAATCGCATCGTGGAAAGCTTTGGGCTTCTACCGATTTAGCAGTTTG	2450
DB	7307	CAGGTAAATTTATATCTTTCTTTTATGTTTTCGGTATTAACCACTAATATCAATTT	7248	DB	TACGATAAGGGGACAAAATCGCATCGTGGAAAGCTTTGGGCTTCTACCGATTTAGCAGTTTG	6228
QY	1431	CTGTGGTTATTAATAAGTGGTTTGGTTCAATTAATGATTAATATCTCTTTCTCT	1490	QY	ATPACATTTTCTTAAGTATCCACCTGAATCATTAATTCGGGAAATAGAGAAAAATTTGACC	2510
DB	7247	CTGTGGTTATTAATAAGTGGTTTGGTTCAATTAATGATTAATATCTCTTTCTCT	7188	DB	ATPACATTTTCTTAAGTATCCACCTGAATCATTAATTCGGGAAATAGAGAAAAATTTGACC	6168
QY	1491	TCCAAATGCTTAATCAATTTTATAAGTTTCAATTTGATATGCTCCTCAATTTTATCT	1550	QY	ATGCTGAAGGCGCCAAATCTGATTTCCACCTGAGATGCATTAATCTAGTAGAAATCTCTCGCT	2570
DB	7187	TCCAAATGCTTAATCAATTTTATAAGTTTCAATTTGATATGCTCCTCAATTTTATCT	7128	DB	ATGCTGAAGGCGCCAAATCTGATTTCCACCTGAGATGCATTAATCTAGTAGAAATCTCTCGCT	6108
QY	1551	AAAGTGAATTTAGAGGCTTACTGCTGCTCTTCTTCAATAGAAATCAATCCTTTTAAA	1610	QY	ATCAAAAATCACTTCCACTCACCGGTTGTCCATTCAGCTGGTGAATCTGCTTTC	2630
DB	7127	AAAGTGAATTTAGAGGCTTACTGCTGCTCTTCTTCAATAGAAATCAATCCTTTTAAA	7068	DB	ATCAAAAATCACTTCCACTCACCGGTTGTCCATTCAGCTGGTGAATCTGCTTTC	6048
QY	1611	AGTCAATTAATGTAACATAAATATATATTTTAAAAATATCCACATTTTCAATTTTC	1670	QY	CTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACC	2690
DB	7067	AGTCAATTAATGTAACATAAATATATATTTTAAAAATATCCACATTTTCAATTTTC	7008	DB	CTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACC	5988
QY	1671	GTTTGTGAACATATGAGTGTCTTAGTTGAAGATTAAGAACCAATTAAGAAATGTTGTC	1730	QY	AAAGAGGCGCCATAAACCAATAGCCTTAACATCATCCCATATTTTATCCAAATATTCGTTTC	2750
DB	7007	GTTTGTGAACATATGAGTGTCTTAGTTGAAGATTAAGAACCAATTAAGAAATGTTGTC	6948	DB	AAAGAGGCGCCATAAACCAATAGCCTTAACATCATCCCATATTTTATCCAAATATTCGTTTC	5928
QY	1731	TTTTCTGTTTTTTTAAAGGATTCAGCGGTAGCGAATAATCCTTTCTTTCTTATCTTGAT	1790	QY	CTTAATTTTATGAGAACAAATCTTCAATCTTCTCTCTAGTCAATTAATTTGCTCCATTCAC	2810
DB	6947	TTTTCTGTTTTTTTAAAGGATTCAGCGGTAGCGAATAATCCTTTCTTTCTTATCTTGAT	6888	DB	CTTAATTTTATGAGAACAAATCTTCAATCTTCTCTCTAGTCAATTAATTTGCTCCATTCAC	5868
QY	1791	AATAAGGGTAACTATGCGGGATAGACTGTAAACATTTCTACGCAATAAAATCCCTTTTCA	1850	QY	TATTTCTCAATTCCTTTTTCAGATAATTTTAGATTGCTTTTCTAAATAGAAATATTTGGAG	2870
DB	6887	AATAAGGGTAACTATGCGGGATAGACTGTAAACATTTCTACGCAATAAAATCCCTTTTCA	6828	DB	TATTTCTCAATTCCTTTTTCAGATAATTTTAGATTGCTTTTCTAAATAGAAATATTTGGAG	5808
QY	1851	TTTTCTAATGTAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAAT	1910	QY	TATTTCTCAATTCCTTTTTCAGATAATTTTAGATTGCTTTTCTAAATAGAAATATTTGGAG	2990
DB	6827	TTTTCTAATGTAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAAT	6768	DB	TATTTCTCAATTCCTTTTTCAGATAATTTTAGATTGCTTTTCTAAATAGAAATATTTGGAG	5688
QY	1911	TCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTTATTCGGTTAATGCGCCATGAC	1970	QY	AGCACCGTTTCTTATTCAGCTATTAATACTCGTCTTCTCTAGCATTAATTTGCTCCATTCAC	2930
DB	6767	TCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTTATTCGGTTAATGCGCCATGAC	6708	DB	AGCACCGTTTCTTATTCAGCTATTAATACTCGTCTTCTCTAGCATTAATTTGCTCCATTCAC	5748

```

Db      5627  |||||CGTCATCATCTGATGAATCAATCGCTCTCTCTGTCATCAAGGTTATTT 5568
Qy      3111  TTATGTAATTTCTTTTAAACAACCAATAGAGATTAAACCTTTTACGGTGTAAACCTTC 3170
Db      5567  TTATGTAATTTCTTTTAAACAACCAATAGAGATTAAACCTTTTACGGTGTAAACCTTC 5508
Qy      3171  CTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTGCATAAAATCGGTATC 3230
Db      5507  CTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTGCATAAAATCGGTATC 5448
Qy      3231  CTTTACAGATATTTTGAGTTTCTGCAATTCGCGAATGTATCCGATTATATTTATT 3290
Db      5447  CTTTACAGATATTTTGAGTTTCTGCAATTCGCGAATGTATCCGATTATATTTATT 5388
Qy      3291  TTTTCGGTATTTTATTAAA 3310
Db      5387  TTTTCGGTGAATCAATTTGAA 5368

RESULT 5
AAN70271/c
ID AAN70271 standard; DNA; 2059 BP.
XX
AC AAN70271;
XX
DT 01-JAN-1980 (first entry)
DE DNA fragment for transformation of lactic acid bacteria.
KW Lactic acid bacterium; antibiotic-resistance gene; food; cheese; dairy;
XX plasmid; vector; ss.
OS Synthetic.
XX
PN EP228726-A.
XX
PD 15-JUL-1987.
XX
PF 20-NOV-1986; 86EP-00202061.
XX
PR 29-NOV-1985; 85NL-00003316.
XX
PA (NEZU-) NED INST ZUIVEL.
XX
PS WPI; 1987-193108/28.
XX
DR DNA for transforming lactic acid bacteria - comprising partial fragment
PT PT which codes for a replicon activity in lactic acid bacteria and an
PT antibiotic-resistance gene.
XX
PS Disclosure; Fig 2; 23pp; English.
XX
CC This sequence contains DNA encoding a replicon activity in lactic acid
CC bacteria and at least 1 antibiotic-resistance gene; either kanamycin-
CC resistance and/or chloramphenicol-resistance. Application of this
CC sequence and a plasmid containing it may lead to the development of
CC lactic acid bacteria which produce polypeptides in a stable and efficient
CC manner. The produced lactic acid bacteria would have application in the
CC food industry. See also AAN70272 and AAN70189
XX
SQ Sequence 2059 BP; 672 A; 424 C; 263 G; 700 T; 0 U; 0 Other;

Query Match 32.1%; Score 1650.4; DB 1; Length 2059;
Best Local Similarity 99.5%; Pred. No. 1.2e-287;
Matches 1677; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

Qy      3291  TTTTCGGTATTTTATTAAAACGTTCTCAAAATCGTTTCTGGGACGTTTTAGCGTTTATT 3350
Db      1685  TTATCGAATTTTATTAAAACGTTCTCAAAATCGTTTCTGAGACGTTTTCGTTTATT 1626
Qy      3351  CGTTTAGTATCGGATATCTTTAAACAGCGGTATTCGTAGCGTAAAGCCCTTGAGC 3410
```

```

Db      1625  CGTTTAGTATCGGATATCTTTAAACAGCGGTTTATCGTAGCTAAAGCCCTTGAGC 1566
Qy      3411  GTAGCGTCTTCCACGGAAGATGTTGCTGTAGATTATGAAGCCGATGACTGAATCA 3470
Db      1565  GTAGCGTCTTCCACGGAAGATGTTGCTGTAGATTATGAAGCCGATGACTGAATCA 1506
Qy      3471  AATAATAAGCGCGATCGCTCTTCTATTTCGGTTGAGGAGGCTCAAGGGAGTTTGAAGGAA 3530
Db      1505  AATAATAAGCGCGATCGCTCTTCTATTTCGGTTGAGGAGGCTCAAGGGAGTTTGAAGGAA 1446
Qy      3531  TGAATTCCTCTCATGGGTTTGATTTAAATAATTCCTTGCATTTTCCGAGCGGTAGCGC 3590
Db      1445  TGAATTCCTCTCATGGGTTTGATTTAAATAATTCCTTGCATTTTCCGAGCGGTAGCGC 1386
Qy      3591  TGG-AAAAATTTTGAATAAAATTTGCAATTTGGAATAATGGGGGAAAGGAGCGAAT 3649
Db      1385  TGGAAAAATTTTGAATAAAATTTGCAATTTGGAATAATGGGGGAAAGGAGCGAAT 1326
Qy      3650  TTTGCTTCGTACTACGACCCCCCATTAAGTCCGAGTCCCAATTTTGTGCCAAAACG 3709
Db      1325  TTTGCTTCGTACTACGACCCCCCATTAAGTCCGAGTCCCAATTTTGTGCCAAAACG 1266
Qy      3710  CTCTATCCCAACTGGCTCAAGGGTTTGAGGGGTTTTTCAATCGCCCAACGAATGCCAACG 3769
Db      1265  CTCTATCCCAACTGGCTCAAGGGTTTGAGGGGTTTTTCAATCGCCCAACGAATGCCAACG 1206
Qy      3770  TTTTCGCCAACGTTTTTATAAATCTATATTAAAGTAGCTTTTATCTGTTGTTTATGATT 3829
Db      1205  TTTTCGCCAACGTTTTTATAAATCTATATTAAAGTAGCTTTTATCTGTTGTTTATGATT 1146
Qy      3830  ACAAAGTGATACACTAATTTTATAAATTTATTTGATGGAGTTTTTTAAATCGTGATTC 3889
Db      1145  ACAAAGTGATACACTAATTTTATAAATTTATTTGATGGAGTTTTTTAAATCGTGATTC 1086
Qy      3890  AGAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAAAAAATTAACAGA 3949
Db      1085  AGAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAAAAAATTAACAGA 1026
Qy      3950  TATGGCGAAAACAAAAGGTTTTTCAAAATCTCGGGTTGCGGCTTAGCTATAGAGAATA 4009
Db      1025  TATGGCGAAAACAAAAGGTTTTTCAAAATCTCGGGTTGCGGCTTAGCTATAGAGAATA 966
Qy      4010  TGAAGAAAGGAATCAGACAAAATAAGCGAAAGCTCGCGTTTTTAGAAGGATACG 4069
Db      965  TGAAGAAAGGAATCAGACAAAATAAGCGAAAGCTCGCGTTTTTAGAAGGATACG 906
Qy      4070  AGTTTCGCTACTGTTTTTGATAAGTAATATATCATGCTATTAAAAATCTAAAGCT 4129
Db      905  AGTTTCGCTACTGTTTTTGATAAGTAATATATCATGCTATT--AAATCTAAAGCT 848
Qy      4130  AGAATTTTGGATTTTATTATATCTGACTCAATTCCTAAATGATTGGAAAGAAATTA 4189
Db      847  AGAATTTTGGATTTTATTATATCTGACTCAATTCCTAAATGATTGGAAAGAAATTA 788
Qy      4190  GAGAGTTTGGCGGTATCTATGGCTGTCAGTCTCTTACAGCATATGACGAAAAAAGAT 4249
Db      787  GAGAGTTTGGCGGTATCTATGGCTGTCAGTCTCTTACAGCATATGACGAAAAAAGAT 728
Qy      4250  AAAGATACATGGAATAGTAGTATGATTACGAAATGGAAGACATATAAAACACACAC 4309
Db      727  AAAGATACATGGAATAGTAGTATGATTACGAAATGGAAGACATATAAAACACACAC 668
Qy      4310  TATCAGCTTATATATATGACGAAATCTGTAAACAATAGAAAGCGTTAGGAACAAGATT 4369
Db      667  TATCAGCTTATATATATGACGAAATCTGTAAACAATAGAAAGCGTTAGGAACAAGATT 608
Qy      4370  AAGCGAAATTTGGGAATAGTTTCAGTTGCTCATGTTGAGATACCTTGATTATATCAAGGT 4429
Db      607  AAGCGAAATTTGGGAATAGTTTCAGTTGCTCATGTTGAGATACCTTGATTATATCAAGGT 548
Qy      4430  TCATATCAATTTTGTACTCATGAATCAAAGACGCTATTTCCTAAGAAATTAACATATAC 4489
Db      547  TCATATGAATTTTGTACTCATGAATCAAAGACGCTATTTCCTAAGAAATTAACATATAC 488
```

QY	4490	GACAAAAAGATATTTTGAACATTAATGATTTTGATATTCACCGCTATATAACACATTGAT	4549
Db	487	GACAAAAAGATATTTTGAACATTAATGATTTTGATATTCACCGCTATATAACACATTGAT	428
QY	4550	GAAAGCCAAAAAGAGATTTGAAGATTTTACTTTTAGATATAGTGGATGACTATAATTTC	4609
Db	427	GAAAGCCAAAAAGAGATTTGAAGATTTTACTTTTAGATATAGTGGATGACTATAATTTC	368
QY	4610	GTAATATCAAAAGATTTAAATGGCTTTTATTCGCCCTTAGGGGAGCGGAGTTTGGAAATTTA	4669
Db	367	GTAATATCAAAAGATTTAAATGGCTTTTATTCGCCCTTAGGGGAGCGGAGTTTGGAAATTTA	308
QY	4670	AATACGAATGATGTAAGAATTTTGTTCACAAACTCTAGCGCTTTAGATATGTTTT	4729
Db	307	AATACGAATGATGTAAGAATTTTGTTCACAAACTCTAGCGCTTTAGATATGTTTT	248
QY	4730	GAGGCAATATCATGTGGATATAGACCAAGTTATGCAAGGTTCTTGTATGCTGAAACG	4789
Db	247	GAGGCAATATCATGTGGATATAGACCAAGTTATGCAAGGTTCTTGTATGCTGAAACG	188
QY	4790	GGGGAATTAATAATGACAAACAAAGAAAGATTTATTTGCTGAAATGAGGAATTAATAA	4849
Db	187	GGGGAATTAATAATGACAAACAAAGAAAGATTTATTTGCTGAAATGAGGAATTAATAA	128
QY	4850	AAGAAATTAAGGACTTTAAAGAGCGTATTGAAGATACAGAGAATGGAAGTTGAATTA	4909
Db	127	AAGAAATTAAGGACTTTAAAGAGCGTATTGAAGATACAGAGAATGGAAGTTGAATTA	68
QY	4910	GTACAAATATAGATTTATTGAGAGAGGAGTTATTGAAATTAATAAAGCCCGCTGACGA	4969
Db	67	GTACAAATATAGATTTATTGAGAGAGGAGTTATTGAAATTAATAAAGCCCGCTGACGA	8
QY	4970	AAGTCG 4975	
Db	7	AAGTCG 2	
RESULT 5			
QY	AAQ48463	AAQ48463 standard; DNA; 3792 BP.	
AC	AAQ48463;		
DT	25-MAR-2003 (revised)		
DT	18-MAR-1994 (first entry)		
DE	Plasmid pG-host4 containing Ts replication system.		
DE	Temperature sensitive replication; antibiotic resistance marker gene;		
KW	site-specific recombination; chromosomal integration; inactivation;		
KW	heterologous gene expression; thermosensitive plasmid; ds.		
OS	Synthetic.		
PN	WO9318164-A1.		
XX	16-SEP-1993.		
XX	12-MAR-1993; 93WO-FR000248.		
PR	13-MAR-1992; 92FR-00003034.		
XX	(INRG ) INRA INST NAT RECH AGRONOMEQUE.		
XX	Gruss A, Maguin E;		
DR	WPI; 1993-303478/38.		
PT	New bacterial plasmid contg. heat sensitive replication system - and		
PT	marker gene, opt. capable of chromosomal integration, used to inactivate		
PT	specific gene or introduce heterologous gene.		
XX			

PS	Example 2; Fig 9; 73pp; French.		
XX	Plasmid pK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two		
CC	antibiotic resistance marker genes was subjected to mutagenesis with		
CC	hydroxylamine. A heat-stable mutant was isolated (coding for a heat-		
CC	sensitive RepA - see AAQ48466 and AAQ48467), cut with ClaI and HpaII and		
CC	the 340bp fragment lacking the Cm resistance gene was ligated to a 445bp		
CC	PvuII fragment of pBluescript SK+ containing a multicloning site, T7 and		
CC	T3 promoters and sequencing primer binding sites. The resulting plasmid		
CC	was designated pV6004 (or pG-host4 - AAQ48463). It is thermosensitive in		
CC	all hosts tested, including E.coli, and must be maintained at 28 deg.C.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
XX			
SQ	Sequence 3792 BP; 1249 A; 552 C; 742 G; 1249 T; 0 U; 0 Other;		
Query Match	31.8%; Score 1636.8; DB 2; Length 3792;		
Best Local Similarity	98.3%; Pred. No. 3.7e-285;		
Matches 1686;	Conservative 0; Mismatches 25; Indels 3; Gaps 3;		
QY	3279 TTTATATTTATTTTCGGTATTTTTTATTAACGCTCTCAAAATCGTTTCGGACGTTT	3338	
Db	1181 TATATATATTTATTTATCGCATTTTTTATTAACGCTCTCAAAATCGTTTCGGACGTTT	1240	
QY	3339 TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTTAAACAGCGGTTATCGTAGCGTAA	3398	
Db	1241 TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTTAAACAGCGGTTATCGTAGCGTAA	1300	
QY	3399 AAGCCCTTGAAGCGTAGCGT-GCTTTGCAGCGAAGATGTTCTGTTAGATTATGAAGCC	3457	
Db	1301 AAGCCCTTGAAGCGTAGCGTGGCTTTGCAGCGAAGATGTTCTGTTAGATTATGAAGCC	1360	
QY	3458 GATGACTGAATGAATAATAAAGCGCAGCGCTCTCTATTTTCGGTTGGAGGCTCAAGG	3517	
Db	1361 GATGACTGAATGAATAATAAAGCGCAGCGCTCTCTATTTTCGGTTGGAGGCTCAAGG	1420	
QY	3518 GAGTTGAGGGAATGAATTCCTCATGGTTTGAATTTTAAATAATGCTGCAATTTTGC	3577	
Db	1421 GAGTATGAGGGAATGAATTCCTCATGGTTTGAATTTTAAATAATGCTGCAATTTTGC	1480	
QY	3578 CGAGCGGTAGCGCTCGAAAAATTTTGAATAAATTTTGAATTTGAAAAAATGGGGGA	3637	
Db	1481 CGAGCGGTAGCGCTCGAAAAATTTTGAATAAATTTTGAATTTGAAAAAATGGGGGA	1540	
QY	3638 AAGGAAGCGAATTTTTCCTCGTACTCGACCCCATTAAGTCCGAGTGCCTATTTT	3697	
Db	1541 AAGGAAGCGAATTTTTCCTCGTACTCGACCCCATTAAGTCCGAGTGCCTATTTT	1600	
QY	3698 GTGCCAAAAACGCTCTATCCCAACTGGCTCAAGGGTTTTCAGTCCCAAC	3757	
Db	1601 GTGCCAAAAACGCTCTATCCCAACTGGCTCAAGGGTTTTCAGTCCCAAC	1660	
QY	3758 GAATCCCAACGTTTTCGCAACGTTTATTAATCTATATTAAGTAGCTTTATGTT	3817	
Db	1661 GAATCCCAACGTTTTCGCAACGTTTATTAATCTATATTAAGTAGCTTTATGTT	1720	
QY	3818 GTTTTATGATTACAAAGTGATACACTAATTTTATAAAATTTATTTGATGGAGTTT	3877	
Db	1721 GTTTTATGATTACAAAGTGATACACTAATTTTATAAAATTTATTTGATGGAGTTT	1780	
QY	3878 AATGGTGAATTCAGATATCGAAAAAAGAGTTATGATTTCTCGACAAAGACAGATAA	3937	
Db	1781 AATGGTGAATTCAGATATCGAAAAAAGAGTTATGATTTCTCGACAAAGACAGATAA	1840	
QY	3938 AAAATTAACAGATATCGCGAAAAACAAAGGTTTTCATAAATCTGCGGCTTAGC	3997	
Db	1841 AAAATTAACAGATATCGCGAAAAACAAAGGTTTTCATAAATCTGCGGCTTAGC	1900	
QY	3998 TATAGAAGATATGCAAGAAAGGAATTCAGAAACAAAAAATTAAGCGAGCTCGGTTT	4057	
Db	1901 TATAGAAGATATGCAAGAAAGGAATTCAGAAACAAAAAATTAAGCGAGCTCGGTTT	1960	
QY	4058 TAGAAGGATACGAGTTTTCGCTACTGTTTGTATTAAGGTAA-TATATCATGGCTATTAA	4116	

Db 1961 TAGAAGGATACGAGTTTCGCTACTGTTTGTGATAAGGTAATATATATCATGGCTATTAA 2020  
Qy 4117 AAATACTAAAGCTAGAAATTTTGGATTTTATATATCTGACTCAATTCCTTAATGATTG 4176  
Db 2021 AAATACTAAAGCTAGAAATTTTGGATTTTATATATCTGACTCAATTCCTTAATGATTG 2080  
Qy 4177 GAAAGAAAATTTAGAGAGTTTGGGCGTATCTATGCTGTCTAGTCTTTACACGATATGGA 4236  
Db 2081 GAAAGAAAATTTAGAGAGTTTGGGCGTATCTATGCTGTCTAGTCTTTACACGATATGGA 2140  
Qy 4237 GAAAGAAAATTTAGAGATACATGATGAATAGTAGTGTATACGAATGAAAGACACTA 4296  
Db 2141 GAAAGAAAATTTAGAGATACATGATGAATAGTAGTGTATACGAATGAAAGACACTA 2200  
Qy 4297 TAAAAAACACACTATCATGTTATATATATTTGACACGAATCTGTGACAAATAGAAAAGCGT 4356  
Db 2201 TAAAAAACACACTATCATGTTATATATATTTGACACGAATCTGTGACAAATAGAAAAGCGT 2260  
Qy 4357 TAGGACACAGATTAGCGGAATTTGGGAATGTTGCTGCTCATGTTGAGATACITGA 4416  
Db 2261 TAGGACACAGATTAGCGGAATTTGGGAATGTTGCTGCTCATGTTGAGATACITGA 2320  
Qy 4417 TTATATCAAGGTTTCATATGATATTTGACTCATGAATCAAGGACGCTATTGCTAAGAA 4476  
Db 2321 TTATATCAAGGTTTCATATGATATTTGACTCATGAATCAAGGACGCTATTGCTAAGAA 2380  
Qy 4477 TAAACATATATACGACAAAAGATATTTTGAACATTAATGATTTTGTATATGACCGCTA 4536  
Db 2381 TAAACATATATACGACAAAAGATATTTTGAACATTAATGATTTTGTATATGACCGCTA 2440  
Qy 4537 TATAACACTGTGATGAAGCCAAAAGAGAAATTTGAAGAAATTTACATTTTGTATATAGTGA 4596  
Db 2441 TATAACACTGTGATGAAGCCAAAAGAGAAATTTGAAGAAATTTACATTTTGTATATAGTGA 2500  
Qy 4597 TGACTATATTTGTTGTAATATACAAAAGATTTTAAATGCTTTTATTCGCTTTAGGGGAGCGGA 4656  
Db 2501 TGACTATATTTGTTGTAATATACAAAAGATTTTAAATGCTTTTATTCGCTTTAGGGGAGCGGA 2560  
Qy 4657 GTTTCGAAATTTAAATACGATGATGATAAGATATTTGTTTCAACAACTCTAGCGCTT 4716  
Db 2561 GTTTCGAAATTTAAATACGATGATGATAAGATATTTGTTTCAACAACTCTAGCGCTT 2620  
Qy 4717 TAGATTATGTTTGGGCAATTTATCAGTGTGGATATAGACCAAGTTATGCAAAAGGTTCT 4776  
Db 2621 TAGATTATGTTTGGGCAATTTATCAGTGTGGATATAGACCAAGTTATGCAAAAGGTTCT 2680  
Qy 4777 TGATCTGAACGGGGAAATTAATGACAAACAAAGAAAAGAGTTATTTGCTGAAAAT 4836  
Db 2681 TGATCTGAACGGGGAAATTAATGACAAACAAAGAAAAGAGTTATTTGCTGAAAAT 2740  
Qy 4837 GAGGAATTAAGAAAAGAAATTAAGGACTTAAAGAGCGCTATTGAAAGATACAGAGAAATG 4896  
Db 2741 GAGGAATTAAGAAAAGAAATTAAGGACTTAAAGAGCGCTATTGAAAGATACAGAGAAATG 2800  
Qy 4897 GAAGTTGAATTAAGTACAAATAGATTTATTTGAGAGAGGATTTATTTGAATAATAAAA 4956  
Db 2801 GAAGTTGAATTAAGTACAAATAGATTTATTTGAGAGAGGATTTATTTGAATAATAAAA 2860  
Qy 4957 GCGCCCTGACGAAAGTCGACCTCGTCTCTTTT 4992  
Db 2861 G-CCCCCTGACGAAAGTCGAGGGGGTTTATTTT 2895

RESULT 7  
ID AAQ48464  
XX AAQ48464 standard; DNA; 5234 BP.  
AC AAQ48464;  
XX  
XX 25-MAR-2003 (revised)  
DT 18-MAR-1994 (first entry)  
XX  
DE Plasmid pG+host5 containing Ts replication system.

XX  
KW Temperature sensitive replication; antibiotic resistance marker gene;  
KW site-specific recombination; chromosomal integration; inactivation;  
KW heterologous gene expression; thermosensitive plasmid; ds.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FH misc\_feature 36..1496  
FT /\*tag= a  
FT /standard\_name= "ORI"  
FT /note= "Origin of replication from pBR322"  
FT 2640..4383  
FT /\*tag= b  
FT /note= "from pGK12 (derived from pWV01)"  
FT 4384..4786  
FT /\*tag= c  
FT /note= "from pUB110"  
FT 4787..5234  
FT /\*tag= d  
FT /note= "from pSK"  
XX  
PN WO9318164-A1.  
PD 16-SEP-1993.  
XX  
XX 12-MAR-1993; 93WO-FR000248.  
XX  
XX 13-MAR-1992; 92PR-00003034.  
XX  
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX  
XX Gruss A, Maguin E;  
XX  
XX WPI; 1993-303478/38.  
XX  
XX New bacterial plasmid contg. heat sensitive replication system - and  
XX marker gene, opt. capable of chromosomal integration, used to inactivate  
XX specific gene or introduce heterologous gene.  
XX  
XX Example 2; Fig 10; 73pp; French.  
XX  
XX Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two  
XX antibiotic resistance marker genes was subjected to mutagenesis with  
XX hydroxylamine. A heat-stable mutant was isolated (coding for a heat-  
XX sensitive RepA - see AAQ48466 and AAQ48467), cut with ClaI and HpaII and  
XX the 3340bp fragment lacking the Cm resistance gene was ligated to a 445bp  
XX T3 promoter and sequencing primer binding sites. The resulting plasmid  
XX was designated pVE6004 (or pG+host4 - AAQ48463). To facilitate cloning in  
XX E.coli, the 1.4kb Aval-AluNI fragment of pBR322 (containing the origin of  
XX replication) was inserted into NsiI-cleaved pG+host4 to give pG+host5  
XX (AAQ48464). The pBR322 ORI allows maintenance of the plasmid in E.coli at  
XX 37 deg.C. while the heat-sensitive ORI allows maintenance at 28 deg.C in  
XX gram-positive bacteria. (Updated on 25-MAR-2003 to correct FN field.)  
SQ Sequence 5234 BP; 1569 A; 973 C; 1133 G; 1559 T; 0 U; 0 Other;

Query Match 31.8%; Score 1635.2; DB 2; Length 5234;  
Best Local Similarity 98.2%; Pred. No. 7.5e-285;  
Matches 1685; Conservative 0; Mismatches 28; Indels 3; Gaps 3;  
Qy 3279 TTTATATTTTATTTTCGGTATTTTATTTTAAACGCTCAAAATCGTTTCTGGAGCTTT 3338  
Db 2623 TATATATATTTATTTATCCGATTTTATTTTAAAGCTCAAAATCGTTTCTGGAGCTTT 2682  
Qy 3339 TAGCGTTTATTTCTGTAGTTATCGGCATAATCGTTTAAACAGCGCTTATCGTAGCTAA 3398  
Db 2683 TAGCGTTTATTTCTGTAGTTATCGGCATAATCGTTTAAACAGCGCTTATCGTAGCTAA 2742  
Qy 3399 AAGCCCTTGAGCGTAGCGT-GCTTTTCAGCAGAGATGTTGTTCTGTAGATTATGAAGCC 3457  
Db 2743 AAGCCCTTGAGCGTAGCGTGGCTTTCAGCAGAGATGTTGTTCTGTAGATTATGAAGCC 2802

QY	3458	GATGACTGAATCAAAATATTAAGCCGACGGTCCCTTCTATTTCGGTTGGAGGAGGCTCAAGG	3517
DB	2803	GATGACTGAATCAAAATATTAAGCCGACGGCCCTTCTATTTCGGTTGGAGGAGGCTCAAGG	2862
QY	3518	GAGTTTGAGGGAATGAAATCCCTCATCGGTTTGATTTTAAAAATTCGTGCAATTTTGC	3577
DB	2863	GAGTATGAGGGAATGAAATCCCTCATCGGTTTGATTTTAAAAATTCGTGCAATTTTGC	2922
QY	3578	CGAGCGGTAGCGCTGGAATAATTTTGAATAAATAATTTGGAATTTGGAATAAATGCGGGGA	3637
DB	2923	CGAGCGGTAGCGCTGGAATAATTTTGAATAAATAATTTGGAATTTGGAATAAATGCGGGGA	2982
QY	3638	AAGGAAGCGAATTTTGTCTCCGTACTACGACCCCCCAATTAAGTTCGCGAGTCCCAATTTT	3697
DB	2983	AAGGAAGCGAATTTTGTCTCCGTACTACGACCCCCCAATTAAGTTCGCGAGTCCCAATTTT	3042
QY	3698	GTGCCAAAAACGCTCTATCCCAACTGCTCAAGGTTTCAGGGGTTTTTCATCGCCCAAC	3757
DB	3043	GTGCCAAAAACGCTCTATCCCAACTGCTCAAGGTTTAAAGGGTTTTTCAATCGCCCAAC	3102
QY	3758	GAAATGCCAACGTTTTTCGCCAACGTTTTTATAAATCTATATTAAAGTAGCTTTTATGTT	3817
DB	3103	GAAATGCCAACGTTTTTCGCCAACGTTTTTATAAATCTATATTAAAGTAGCTTTTATGTT	3162
QY	3818	GTTTTTATGATTTACAAAGTGATACACTAAATTTTATAAAAAATTAATTTGATTGGAGTTTTT	3877
DB	3163	GTTTTTATGATTTACAAAGTGATACACTAACTTTATATAAATTAATTTGATTGGAGTTTTT	3222
QY	3878	AATGCTGATTTTCAGNATCGAAAAAAGAGTTATGATTTCTCTGACAAAAAGCAAGATAA	3937
DB	3223	AATGCTGATTTTCAGNATCGAAAAAAGAGTTATGATTTCTCTGACAAAAAGCAAGATAA	3282
QY	3938	AAAAATTAACAGATATGGCGAAAACAAAAAGGTTTTTCAAAATCTCGGGTTGGCGCGTTAGC	3997
DB	3283	AAAAATTAACAGATATGGCGAAAACAAAAAGGTTTTTCAAAATCTCGGGTTGGCGCGTTAGC	3342
QY	3998	TATAGAGAAATATCCAGNAGGAAATCAGAACAAAAAATAAGCGAAGCTCGCGTTTTT	4057
DB	3343	TATAGAGAAATATCCAGNAGGAAATCAGAACAAAAAATAAGCGAAGCTCGCGTTTTT	3402
QY	4058	TAGAAGGATACGAGTTTTCGCTACTTGTTTTTGATAAGGTAA-TATATCATCGGCTATTAA	4116
DB	3403	TAGAAGGATACGAGTTTTCGCTACTTGTTTTTGATAAGGTAAATATATCATCGGCTATTAA	3462
QY	4117	AAATACTAAAGCTAGAAATTTTGGATTTTTTATTATATCTGACTCAATTCCTCTAATGATTG	4176
DB	3463	AAATACTAAAGCTAGAAATTTTGGATTTTTTATTATATCTGACTCAATTCCTCTAATGATTG	3522
QY	4177	GAAAGAAAAATTAGAGAGTTTGGGCGTATCTATGCGCTGCAGTCTCTTTTACACGATFGGA	4236
DB	3523	GAAAGAAAAATTAGAGAGTTTGGGCGTATCTATGCGCTGCAGTCTCTTTTACACGATFGGA	3582
QY	4237	CGAAAAAAAACATAAAGATACATGGAATAGTAGTGATGTATACGAATGGAAGCACTA	4296
DB	3583	CGAAAAAAAACATAAAGATACATGGAATAGTAGTAATATTATACAAATGGAAGCACTA	3642
QY	4297	TAAAAAACCACTATCACGTTATATATTTGACAGAAATCCTGTACAAATAGNAGCGT	4356
DB	3643	TAAAAAACCACTATCACGTTATATATTTGACAGAAATCCTGTACAAATAGNAGCGT	3702
QY	4357	TAGGAACAAGATTAAAGCGAAAAATTTGGGGAATAGTTTCAGTTCCTCATGTTCAGATACCTGA	4416
DB	3703	TAGGAACAAGATTAAAGCGAAAAATTTGGGGAATAGTTTCAGTTCCTCATGTTCAGATACCTGA	3762
QY	4417	TTATATCAAGGTTTCATATCAATATTTGACTCATAGTAACGAGCCGCTATTCCTAGAA	4476
DB	3763	TTATATCAAGGTTTCATATCAATATTTGACTCATAGTAACGAGCCGCTATTCCTAGAA	3822
QY	4477	TAAACATATATACGACAAAAAAGATATTTGAAACATTAATGATTTTGATTTGACCGCTA	4536
DB	3823	TAAACATATATACGACAAAAAAGATATTTGAAACATTAATGATTTTGATTTGACCGCTA	3882

QY	4537	TATTAACACTTTGATGAAGACCCAAAAAGAGAAATTTGAAGAATTTTACTTTTATAGATATAGTGGA	4599
Db	3883	TATAACACTTTGATGAAGACCCAAAAAGAGAAATTTGAAGAATTTTACTTTTATAGATATAGTGGA	3942
QY	4597	TGACTATAATTTTGGTAAATACAAAGAAATTTAATGGCTTTTATTTCGCCTTTAGGGGAGCGGA	4656
Db	3943	TGACTATAATTTTGGTAAATACAAAGAAATTTAATGGCTTTTATTTCGCCTTTAGGGGAGCGGA	4002
QY	4657	GTTTGGAAATTTTAAATACGATGATGTAAGAAGATTTGTTTCAACAAACTCTAGCGCCTT	4716
Db	4003	GTTTGGAAATTTTAAATACGATGATGTAAGAAGATTTGTTTCAACAAACTCTAGCGCCTT	4062
QY	4717	TAGATTATGTTTTCAGGGGCAATTTATCAGTGTGGATATAGACAAAGTTATGCAAAAGTTCT	4776
Db	4063	TAGATTATGTTTTCAGGGGCAATTTATCAGTGTGGATATAGACAAAGTTATGCAAAAGTTCT	4122
QY	4777	TGATGCTGAACCGGGGAAATTAATATGACAAACAAAGAAAAGAGATTTTCTGTAAGAT	4836
Db	4123	TGATGCTGAACCGGGGAAATTAATATGACAAACAAAGAAAAGAGATTTTCTGTAAGAT	4182
QY	4837	GAGGAATTTAAAAAAGAAAATTTAAGGACTTTAAAAGAGCGCTATTCAAAGATACAGAGAAATG	4896
Db	4183	GAGGAATTTAAAAAAGAAAATTTAAGGACTTTAAAAGAGCGCTATTCAAAGATACAGAGAAATG	4242
QY	4897	GAAGTTGAATTAAGTACAACTAGATTTATTGAGAGGAGGATTTATTGCAATATAATAAAA	4956
Db	4243	GAAGTTGAATTAAGTACAACTAGATTTATTGAGAGGAGGATTTATTGCAATATAATAAAA	4302
QY	4957	GCCTCCCTGACGAAAGTCGCGACTTCGTTCTTTT 4992	
Db	4303	G-CCCCCTGACGAAAGTCGAGGGGGTTTATT 4337	

RESULT 8

AAQ48465

ID AAQ48465 standard; DNA; 6722 BP.

XX AC AAQ48465;

XX DT 25-MAR-2003 (revised)

XX DT 18-MAR-1994 (first entry)

XX DE Plasmid pG-host6 containing Ts replication system.

XX DE Temperature sensitive replication; antibiotic resistance marker gene;

XX KW site-specific recombination; chromosomal integration; inactivation;

XX KW heterologous gene expression; thermosensitive plasmid; ds.

XX OS Synthetic.

XX PN WO9318164-A1.

XX PD 16-SEP-1993.

XX PF 12-MAR-1993; 93WO-FR000248.

XX PR 13-MAR-1992; 92FR-00003034.

XX PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.

XX PI Gruss A, Maguin B;

XX DR WPI; 1993-303478/38.

XX PT New bacterial contg. heat sensitive replication system - and

XX PT marker gene, opt. capable of chromosomal integration, used to inactivate

XX PT specific gene or introduce heterologous gene.

XX PS Disclosure; Fig 11; 73pp; French.

XX CC Plasmid pGK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two

XX CC antibiotic resistance marker genes was subjected to mutagenesis with

XX CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-



CC sensitive RepA - see AA048466 and AA048467, cut with ClaI and HpaII and  
CC the 3340bp fragment lacking the Cm resistance gene was ligated to a 445bp  
CC PvuII fragment of pBluescript SK+ containing a multicloning site, 77 and  
CC T3 promoters and sequencing primer binding sites. The resulting plasmid  
CC was designated pV6004 (for pG+host4 - AA048463). To facilitate cloning in  
CC E.coli, the Aval-EcoRI fragment of pBR322 (containing the origin of  
CC replication and the ampicillin resistance gene) was inserted into NsiI-  
CC cleaved, blunt-ended pG+host4 to give pG+host6 (AA048466). The pBR322 ORI  
CC allows maintenance of the plasmid in E.coli at 37 deg.C while the heat-  
CC sensitive ORI allows maintenance at 28 deg.C in gram-positive bacteria.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX

SQ Sequence 6722 BP; 1961 A; 1278 C; 1503 G; 1980 T; 0 U; 0 Other;

Query Match 31.8%; Score 1635.2; DB 2; Length 6722;

Best Local Similarity 98.2%; Pred. No. 7.7e-285;

Matches 1685; Conservative 0; Mismatches 28; Indels 3; Gaps 3;

3279 TTTATATTTATTTTCGGTATTTTATTAACAGCTCTCAAAATCGTTCTCGAGCGTTT 3338

4111 TATATATATTTATTCGGATTTTATTAACAGCTCTCAAAATCGTTCTCGAGCGTTT 4170

3339 TAGCGTTATTTTCGGTATTTTATTAACAGCTCTCAAAATCGTTCTCGAGCGTTT 3398

4171 TAGCGTTATTTTCGGTATTTTATTAACAGCTCTCAAAATCGTTCTCGAGCGTTT 4230

3399 AAGCCCTTGAGCGTAGCGT - GCTTTTCAGCGAAGATGTTTCTGTAGATTATGAAGCC 3457

4231 AAGCCCTTGAGCGTAGCGTGGCTTTTCAGCGAAGATGTTTCTGTAGATTATGAAGCC 4290

3458 GATGACTGAATGAATATAAGCGAGCGTCTCTTATTCGGTTGGAGGAGGCTCAAGG 3517

4291 GATGACTGAATGAATATAAGCGAGCGTCTCTTATTCGGTTGGAGGAGGCTCAAGG 4350

3518 GAGTTTGAGGGAATGAATATTCCTCATGCGTTTGAATTTTAAATATGCTTGCAATTTGC 3577

4351 GAGTATGAGGGAATGAATATTCCTCATGCGTTTGAATTTTAAATATGCTTGCAATTTGC 4410

3578 CGAGCGGTAGCGTGAATAATTTTGAATAAATTTGGAATTTGGAATAAATGCGGGA 3637

4411 CGAGCGGTAGCGTGAATAATTTTGAATAAATTTGGAATTTGGAATAAATGCGGGA 4470

3638 AAGGAAGCGAATTTTTCCTTCCTGCTACGACCCCGATTAAGTCCGAGTGCCTATTTT 3697

4471 AAGGAAGCGAATTTTTCCTTCCTGCTACGACCCCGATTAAGTCCGAGTGCCTATTTT 4530

3698 GTGCCAAAACGCTCATCCCACTGGCTCAAGGGTTTTCAGGGTTTTCAGTCCCAAC 3757

4531 GTGCCAAAACGCTCATCCCACTGGCTCAAGGGTTTTCAGGGTTTTCAGTCCCAAC 4590

3758 GAATCGCCAACTTTTCGCGACAGTTTATTAATCTATATTTAAGTAGCTTTATTTGTT 3817

4591 GAATCGCCAACTTTTCGCGACAGTTTATTAATCTATATTTAAGTAGCTTTATTTGTT 4650

3818 GTTTTATGATTACAAAGTGATACACTAATTTTATAAATATTTTGAATGGAGTTTATTA 3877

4651 GTTTTATGATTACAAAGTGATACACTAATTTTATAAATATTTTGAATGGAGTTTATTA 4710

3878 AATGTTGATTTCAGAAATCGAAATAAAGAGTTATGTTTCTCTGCAAAAGAGCAAGATAA 3937

4711 AATGTTGATTTCAGAAATCGAAATAAAGAGTTATGTTTCTCTGCAAAAGAGCAAGATAA 4770

3938 AAAATTAACAGATATGCGGGAACAAAGGTTTTCAGAAATCGCGGTTGCGGCTAGC 3997

4771 AAAATTAACAGATATGCGGGAACAAAGGTTTTCAGAAATCGCGGTTGCGGCTAGC 4830

3998 TATAGAAGATATGCAAGAAAGGAATACAGAACAAAAAATAAGCGAAGCTCGGTTTT 4057

4831 TATAGAAGATATGCAAGAAAGGAATACAGAACAAAAAATAAGCGAAGCTCGGTTTT 4890

4058 TAGAGGATACAGATTTTCGCTACTGTTTGTATGATGAGTAA - TATATCATGGCTATTAA 4116

4891 TAGAGGATACAGATTTTCGCTACTGTTTGTATGATGAGTAA - TATATCATGGCTATTAA 4950

## RESULT 9

AAN91856/c

ID AAN91856 standard; DNA; 7336 BP.

XX

AAN91856;

AC

XX

25-MAR-2003 (revised)

DT

31-OCT-2002 (revised)

DT

23-MAR-1990 (first entry)

XX

Plasmid pBHA-1.

DE

QY 4117 AATACATAAAGCTAGAAATTTTCGATTTTATATATATCCCTGACTCAATTCCTATGATTC 4176  
DB 4951 AATACATAAAGCTAGAAATTTTCGATTTTATATATATCCCTGACTCAATTCCTATGATTC 5010  
QY 4177 GAAAGAAAAATTAGAGAGTTTGGCGGTATCTATGGCTGTGCTGCTTTTACACGATATGGA 4236  
DB 5011 GAAAGAAAAATTAGAGAGTTTGGCGGTATCTATGGCTGTGCTGCTTTTACACGATATGGA 5070  
QY 4237 CGAAAAAAGATATAAGATACATGATAGTAGTGTATATAGGAATCGGAAGCACTA 4296  
DB 5071 CGAAAAAAGATATAAGATACATGATAGTAGTGTATATAGGAATCGGAAGCACTA 5130  
QY 4297 TAAAAAACCCACATATCACGTTTATATATATGACGAAAAATCCCTGTAAACAATAGAAAGCGT 4356  
DB 5131 TAAAAAACCCACATATCACGTTTATATATATGACGAAAAATCCCTGTAAACAATAGAAAGCGT 5190  
QY 4357 TAGGAACAAGATTAAAGCGAAAAATGGGGAATAGTTTCAGTTGCTGCTCATGTTGAGATACCTGA 4416  
DB 5191 TAGGAACAAGATTAAAGCGAAAAATGGGGAATAGTTTCAGTTGCTGCTCATGTTGAGATACCTGA 5250  
QY 4417 TTATATCAAAAGGTTTCATATGAAATATTTGACTCATGAATCAAGAGCGCTATTCCTAAGAA 4476  
DB 5251 TTATATCAAAAGGTTTCATATGAAATATTTGACTCATGAATCAAGAGCGCTATTCCTAAGAA 5310  
QY 4477 TAAACATATATACGACAAAAAAGATATTTTGAACATTAATGATTTTGAATATGACCGCTA 4536  
DB 5311 TAAACATATATACGACAAAAAAGATATTTTGAACATTAATGATTTTGAATATGACCGCTA 5370  
QY 4537 TATAACACTTTGATCAAGCCAAAGAGAGAAATGGAAGATTTTACTTTTAGATATAGTGA 4596  
DB 5371 TATAACACTTTGATCAAGCCAAAGAGAGAAATGGAAGATTTTACTTTTAGATATAGTGA 5430  
QY 4597 TGACTATAATTTGGTAAATACAAAAAGATTTTAAATGGCTTTTATTCGCCCTTAGGGGAGCGGA 4656  
DB 5431 TGACTATAATTTGGTAAATACAAAAAGATTTTAAATGGCTTTTATTCGCCCTTAGGGGAGCGGA 5490  
QY 4657 GTTTGGAATTTTAAATACGAATGATGTAAAGATATTTGTTTAAACAACTCTAGCGGCTT 4716  
DB 5491 GTTTGGAATTTTAAATACGAATGATGTAAAGATATTTGTTTAAACAACTCTAGCGGCTT 5550  
QY 4717 TAGATATGTTTGGAGGCAATTTATCACTGTGATATAGAGCAAGTTATGCTCAAAAGGTTCT 4776  
DB 5551 TAGATATGTTTGGAGGCAATTTATCACTGTGATATAGAGCAAGTTATGCTCAAAAGGTTCT 5610  
QY 4777 TGATGTGAAAACGCGGGGAAATATAATGACAAAACAAAGAAAAAGAGTTATTTGCTGAAAAAT 4836  
DB 5611 TGATGTGAAAACGCGGGGAAATATAATGACAAAACAAAGAAAAAGAGTTATTTGCTGAAAAAT 5670  
QY 4837 GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGGTATTTGAAAGATACAGAGAAATG 4896  
DB 5671 GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGGTATTTGAAAGATACAGAGAAATG 5730  
QY 4897 GAAAGTTGAATTAAGTACAAATAGATTTTATTCAGAGGAGGAGGATTTTGAATATAAATAA 4956  
DB 5731 GAAAGTTGAATTAAGTACAAATAGATTTTATTCAGAGGAGGAGGATTTTGAATATAAATAA 5790  
QY 4957 GCGCCCTCTGACGAAGTCGCGACTTCGTTCTTTTTT 4992  
DB 5791 G-CGCCCTGACGAAGTCGAAGGCGGTTTTTATTTT 5825



2618 TGAACCTGCTTCTCTCTGACATGACACATCATCTCAATATCGAATAGGCCCAT 2677  
 5527 TGAACCTGCTTCTCTCTGACATGACACATCATCTCAATATCGAATAGGCCCAT 5468  
 2678 CAGTCTGACGACAG 2737  
 5467 CAGTCTGACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5408  
 2738 CCAATATTCGTTCTCTCTCTGACATGACACATCATCTCAATATCGAATAG 2797  
 5407 CCAATATTCGTTCTCTCTCTGACATGACACATCATCTCAATATCGAATAG 5348  
 2798 TTGGTCCATCTCAATCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2857  
 5347 TTGGTCCATCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5288  
 2858 AGAATATTCGTTCT 2917  
 5287 AGAATATTCGTTCT 5228  
 2918 TTCAATCT 2977  
 5227 TTCAATCT 5168  
 2978 ACTACT 3037  
 5167 ACTACT 5108  
 3038 CATCT 3097  
 5107 CATCT 5048  
 3098 CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCCATGAGAGATTAACCTTTTAC 3157  
 5047 CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCCATGAGAGATTAACCTTTTAC 4988  
 3158 GGTGTAACCT 3217  
 4987 GGTGTAACCT 4928  
 3218 TAAATCTCGTATCT 3277  
 4927 TAAATCTCGTATCT 4868  
 3278 ATTATATTTATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTT 3318  
 4867 ATTATATTTATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTT 4827

RESULT 10

ART32227

ID AAT32227 standard; DNA; 8808 BP.

XX

XX

AC AAT32227;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

30-JAN-1995; 95US-00380521.  
 (DUPO ) DU PONT DE NEMOURS & CO E I.  
 Jackson EN, Liu G;  
 WPI; 1996-371435/37.  
 Thermosable enzymes e.g. xylanase and beta-glucosidase - are produced  
 extracellularly, in absence of generally required signal peptide, and in  
 high yields from recombinant bacteria.  
 Example 1; Page 26-31; 48pp; English.  
 Plasmid pBE146 (AAT32227) contains the Bacillus neutral protease gene  
 promoter (nprp). The large NdeI-KpnI fragment of pBE146 was ligated to  
 the 250 bp NdeI-KpnI fragment of plasmid pBE105 (contg. the caldocellum  
 saccharolyticum xylanase gene) and used to construct plasmid pBE145, in  
 which the xynA gene was controlled by nprp. Bacillus subtilis  
 transformants produced extracellular xylanase, despite the lack of a  
 signal peptide. The recombinant enzyme was produced in a yield of 1110  
 U/ml supernatant; this compared with 0 U/ml using a vector (pBE60)  
 control, and 272 U/ml using pBE119 (see also AAT32225-26) in which the  
 xynA gene was controlled by the alkaline protease gene promoter  
 Sequence 8808 BP; 2573 A; 1971 C; 1749 G; 2515 T; 0 U; 0 Other;  
 Query Match 29.0%; Score 1493; DB 2; Length 8808;  
 Best Local Similarity 98.1%; Pred. No. 3.2e-259;  
 Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
 QY 1778 TTCTATCTTGAATAAAGGTAACATATTCGCGGATAGACTGTAACATCTCTCAGCATTA 1837  
 DB 6191 TACTCTTTTAAATATCCCGACTGGCAATGCGGGATAGACTGTAACATCTCTCAGCATTA 6250  
 QY 1838 AAATCCCTTTTCATTTCTTAATGTAATCTATTACCTTATTATTATTGTAATTCGTCAT 1897  
 DB 6251 AAATCCCTTTTCATTTCTTAATGTAATCTATTACCTTATTATTATTGTAATTCGTCAT 6310  
 QY 1898 AATTAACTCTTTTCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTATTCGGTT 1957  
 DB 6311 AATTAACTCTTTTCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTATTCGGTT 6370  
 QY 1958 AATCGCGCATGACAGCCATGATTAATTAATTAATAGGAGAGTTAATAAATACGTAACC 2017  
 DB 6371 AATCGCGCATGACAGCCATGATTAATTAATTAATAGGAGAGTTAATAAATACGTAACC 6430  
 QY 2018 AACATGATTAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2077  
 DB 6431 AACATGATTAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6490  
 QY 2078 CTATATATCGGTGCTGCTTCTGTCACCTCTGTAATCCCATTTCCAGAAATTTCTTAGCGAAT 2137  
 DB 6491 CTATATATCGGTGCTGCTTCTGTCACCTCTGTAATCCCATTTCCAGAAATTTCTTAGCGAAT 6550  
 QY 2138 CCAGAGTTTCTCAGAGTCGGAAGTTGACAGACATTAACGACTGCGACAGATGCTCAT 2197  
 DB 6551 CCAGAGTTTCTCAGAGTCGGAAGTTGACAGACATTAACGACTGCGACAGATGCTCAT 6610  
 QY 2198 AACCTGAAGGAGATCTGATTGCTTAATCTGCTTCAAGTTAAGACCGAGCGCTCGTCGAT 2257  
 DB 6611 AACCTGAAGGAGATCTGATTGCTTAACTGCTTCAAGTTAAGACCGAGCGCTCGTCGAT 6670  
 QY 2258 AACAGATGCGATGATGACAGACCAATCAACATGCGACCTGCCATTTGCTACGTACATCA 2317  
 DB 6671 AACAGATGCGATGATGACAGACCAATCAACATGCGACCTGCCATTTGCTACGTACATCA 6730  
 QY 2318 AGGATGTTAGAAATGTTGTCGGTCTCTGCTGACAGCAATATTACGCATTTGCTGTCATATT 2377  
 DB 6731 AGGATGTTAGAAATGTTGTCGGTCTCTGCTGACAGCAATATTACGCATTTGCTGTCATATT 6790  
 QY 2378 CAAACAGCTCTTCTTACGATTAAGGCGACAAATCGCATCTGCGAATTTGGGCTTCTTACCG 2437

6791 CAACAGCTCTTCTACGATAGGCGACAAATCGCATCGTGGAAGCTTTGGCGTTCTTACCG 6850  
 2438 ATTTAGCAGTTGTATACACTTCTCTAGTATCCACCTGATCATATAATCGGCAAAATAG 2497  
 6851 ATTTAGCAGTTGTATACACTTCTCTAGTATCCACCTGATCATATAATCGGCAAAATAG 6910  
 2498 AGAAAAATGACCATGTGTAGCGGCGCAATCTGATTCACCTGAGATGCATAATCTAGTA 2557  
 6911 AGAAAAATGACCATGTGTAGCGGCGCAATCTGATTCACCTGAGATGCATAATCTAGTA 6970  
 2558 GAATCTCTCGCTATCAAAATCACTTCCAGCTTCCAGCTACCGGTTGTCATTATGCGC 2617  
 6971 GAATCTCTCGCTATCAAAATCACTTCCAGCTTCCAGCTACCGGTTGTCATTATGCGC 7030  
 2618 TGAATCTCTGCTCTCTGTTGACATGACACATCATCTCAATATCGAATAGGCGCCAT 2677  
 7031 TGAATCTCTGCTCTCTGTTGACATGACACATCATCTCAATATCGAATAGGCGCCAT 7090  
 2678 CAGTCTGACGACCAAGAGCGCATAAACACCAATAGCCTTAACATCATCCCATATTTAT 2737  
 7091 CAGTCTGACGACCAAGAGCGCATAAACACCAATAGCCTTAACATCATCCCATATTTAT 7150  
 2738 CCAATATTCGTTCTTAAATTCATGAACAATCTTCAATCTTCTCTAGTCAATATTA 2797  
 7151 CCAATATTCGTTCTTAAATTCATGAACAATCTTCAATCTTCTCTAGTCAATATTA 7210  
 2798 TTGGTCCATCTACTATCTCATCTCCCTTTTCCAGATAATTTAGATTTGCTTTCTAAATA 2857  
 7211 TTGGTCCATCTACTATCTCATCTCCCTTTTCCAGATAATTTAGATTTGCTTTCTAAATA 7270  
 2858 AGAATATTTGGAGACCGCTTCTATTCAGCTATTAATCTCTCTCTCTCTCTCTCTCTCT 2917  
 7271 AGAATATTTGGAGACCGCTTCTATTCAGCTATTAATCTCTCTCTCTCTCTCTCTCTCT 7330  
 2918 TTCAATCTCTTTTAAACAAATATAGCATCTAATCTTCAACAACTGCGCGTTTGTGA 2977  
 7331 TTCAATCTCTTTTAAACAAATATAGCATCTAATCTTCAACAACTGCGCGTTTGTGA 7390  
 2978 ACTACTCTTTTAAACAAATATTTTCCGTTCCCAATCCACATTCGCAATATAGAAATC 3037  
 7391 ACTACTCTTTTAAACAAATATTTTCCGTTCCCAATCCCAATTCGCAATATAGAAATC 7450  
 3038 CATCTTCATCGGCTTTTTCGTCATCATCTGATGATCAATCAAAATCGCTTCTCTGTGCAT 3097  
 7451 CATCTTCATCGGCTTTTTCGTCATCATCTGATGATCAATCAAAATCGCTTCTCTGTGCAT 7510  
 3098 CAAGGTTAAATTTTATGATTTCTTTTAAACAAACCAATAGAGATTAACCTTTTAC 3157  
 7511 CAAGGTTAAATTTTATGATTTCTTTTAAACAAACCAATAGAGATTAACCTTTTAC 7570  
 3158 GGTGTAAACCTTCTCCCAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCA 3217  
 7571 GGTGTAAACCTTCTCCCAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCA 7630  
 3218 TAAATCCGATATCTTTTACAGATATTTTGCAGTTTCTCAATTCGCGATTTGATATCCG 3277  
 7631 TAAATCCGATATCTTTTACAGATATTTTGCAGTTTCTCAATTCGCGATTTGATATCCG 7690  
 3278 ATTATATTTATTTTTCGTTATTTTATTAAGAGTCTCA 3318  
 7691 ATTATATTTATTTTTCGTTATTTTATTAAGAGTCTCA 7731

RESULT 11

AAT18951  
 ID AAT18951 standard; DNA; 9144 BP.

XX  
 AC AAT18951;

XX  
 DT 17-JAN-1997 (first entry)

XX  
 DE Plasmid pBE346.

KW Spider; dragline protein; variant; monomer; polymer; circular; fibre forming region; Spidroin 1; Nepheila clavipes; Dp1; mimic;  
 KW DP-1A analogue; fibre; high tensile strength; elasticity; clothing; rope;  
 KW surgical suture; implant; reinforcement; film; coating; ss.  
 XX Synthetic.  
 OS  
 XX MO9429450-A2.  
 PN  
 XX 22-DEC-1994.  
 PD  
 XX 15-JUN-1994; 94WO-US006689.  
 PF  
 XX 15-JUN-1993; 93US-00077600.  
 PR  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PA  
 XX Fahnstock SR;  
 XX WPI; 1995-036479/05.  
 DR  
 XX  
 PT New synthetic variants of spider dragline protein - for making fibres useful as clothing, surgical silk, plastic reinforcement etc., also related DNA, vectors and transformed cells.  
 PT  
 XX Example 6; Fig 14; 168pp; English.  
 PS  
 XX This sequence represents the complete nucleotide sequence of the plasmid pBE346. This plasmid was used in the construction of the vector pF191 which was used to express synthetic spider dragline variants, DP-1 in B. subtilis. pBE346 comprises a replication origins that confer autonomous replication in both E. coli and B. subtilis, as well as antibiotic resistance makers selectable in E. coli (ampicillin) and B. subtilis (kanamycin). The plasmid also contains the lvs promoter and secretion signal operably linked to a staphylococcal protein A gene. The protein A gene is flanked by an EcoRV site at its proximal end, separating it from the lvs signal sequence, and a BamHI site at its distal end. The protein A gene is replaced by the DP-1 gene. The final plasmid pF191 was used to transform competent cells of B. subtilis BE3010 (trp lys apr apr sacB) Kanamycin resistant transformants of DP-1 is a variant based the fibre forming regions of spider dragline protein, esp. the natural protein 1 (Spidroin 1) from Nepheila clavipes. The proteins may be used to produce fibres of high tensile strength and elasticity, suitable for clothing, rope, surgical sutures, biomaterials for implants, plastic reinforcements, films, coatings, etc

XX Sequence 9144 BP; 2656 A; 2150 C; 1790 G; 2548 T; 0 U; 0 Other;

Query Match 29.0%; Score 1493; DB 2; Length 9144;  
 Best Local Similarity 98.1%; Pred. No. 3.2e-259;  
 Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1778 TTCTATCTTGATAATAAGGTAAGTAACTATTCGCGGATAGACTGTAACATCTTCACGCATA 1837  
 DB 6526 TACTCTTTTAAATATCCCGACTGGCAATGCGCGGATAGACTGTAACATCTTCACGCATA 6585  
 QY 1838 AAATCCCTTTTCAATTTCTAATGTAATCTATTAACCTTATTAATTAATTCGCTCAT 1897  
 DB 6586 AAATCCCTTTTCAATTTCTAATGTAATCTATTAACCTTATTAATTAATTCGCTCAT 6645  
 QY 1898 AATTAATCTTTTCTTATTAACGAAATGGCCGATTAAGCACACCTTTATTCGGTT 1957  
 DB 6646 AATTAATCTTTTCTTATTAACGAAATGGCCGATTAAGCACACCTTTATTCGGTT 6705  
 QY 1958 AATGCCCATGACAGCCATGATAATTAATACTAGGAGAGTAAATAATACGTAAACC 2017  
 DB 6706 AATGCCCATGACAGCCATGATAATTAATACTAGGAGAGTAAATAATACGTAAACC 6765  
 QY 2018 AACATGATTAACAATATTATAGAGTCAATCGTTCAAAATGGTATGCGTTTTCACATCCA 2077  
 DB 6766 AACATGATTAACAATATTATAGAGTCAATCGTTCAAAATGGTATGCGTTTTCACATCCA 6825

QY	2078	CTATATACCGTGTGCGTTCCTGTCACATCTCCGAAATCCCATTTCCAGAAATTTCTCTACGCAATT	2137
Db	6826	CTATATACCGTGTGCGTTCCTGTCACATCTCCGAAATCCCATTTCCAGAAATTTCTCTACGCAATT	6885
QY	2138	CCAGAGTTTCTCAGAGTCGGAAGTTGACGACACATTTACGAACTGGCACAGATGTCAT	2197
Db	6886	CCAGAGTTTCTCAGAGTCGGAAGTTGACGACACATTTACGAACTGGCACAGATGTCAT	6945
QY	2198	AACCTGAAGGAAGATCTGATTGCTTAACCTGCTTACGTTAAGACCGAAGCGCTCGTCGTAAT	2257
Db	6946	AACCTGAAGGAAGATCTGATTGCTTAACCTGCTTACGTTAAGACCGAAGCGCTCGTCGTAAT	7005
QY	2258	AA CAGATGCGATGATGCAGACCAATCAACATGSCACTGCGCATTTGCTACTGCTACAGTCA	2317
Db	7006	AA CAGATGCGATGATGCAGACCAATCAACATGSCACTGCGCATTTGCTACTGCTACAGTCA	7065
QY	2318	AGGATGTTAGAAATGTTGTGCGTCTTTGCGACAGAAATATTACGCCATTTTGCCTGCCATATT	2377
Db	7066	AGGATGTTAGAAATGTTGTGCGTCTTTGCGACAGAAATATTACGCCATTTTGCCTGCCATATT	7125
QY	2378	CAAA CAGCTCTTCTACGATTAAGGCGACAAATCGCATCGTGGACGTTTGGGCTTCTACCG	2437
Db	7126	CAAA CAGCTCTTCTACGATTAAGGCGACAAATCGCATCGTGGACGTTTGGGCTTCTACCG	7185
QY	2438	ATTTAGCAGTTTGATACACATTTTCTCTAAGTATCCACCTGGAATCATTAATCGGCAAAATAG	2497
Db	7186	ATTTAGCAGTTTGATACACATTTTCTCTAAGTATCCACCTGGAATCATTAATCGGCAAAATAG	7245
QY	2498	AGAAAAATTGACCATGTGTGAAGCGGCGAAATCTGATTCACCTGAGATGCATATCTAGTA	2557
Db	7246	AGAAAAATTGACCATGTGTGAAGCGGCGAAATCTGATTCACCTGAGATGCATATCTAGTA	7305
QY	2558	GAATCTCTTCGCTATCAAAATTTCACTTCCACCTTCCACTCACCGTTGTGCCATTCATGGC	2617
Db	7306	GAATCTCTTCGCTATCAAAATTTCACTTCCACCTTCCACTCACCGTTGTGCCATTCATGGC	7365
QY	2618	TGAATCTGCTTCTCTGTGTGACATGACACACATCATCTCAATTCGAAATAGGGCCCAT	2677
Db	7366	TGAATCTGCTTCTCTGTGTGACATGACACACATCATCTCAATTCGAAATAGGGCCCAT	7425
QY	2678	CAGTCTGACGACCAAGAGAGCCATAAACCAATAGCCTTAACATCATCATCCCCATATTAT	2737
Db	7426	CAGTCTGACGACCAAGAGAGCCATAAACCAATAGCCTTAACATCATCCCCATATTAT	7485
QY	2738	CCAATATTGCTTCCCTTAATTTTCATGAACAATCTTTCATTTCTTCTCTAGTCAATTA	2797
Db	7486	CCAATATTGCTTCCCTTAATTTTCATGAACAATCTTTCATTTCTTCTCTAGTCAATTA	7545
QY	2798	TTGGTCCATTCACATATCTCATTCGCCCTTTGAGATAATTTAGATTTGCTTTTCTTAATA	2857
Db	7546	TTGGTCCATTCACATATCTCATTCGCCCTTTGAGATAATTTAGATTTGCTTTTCTTAATA	7605
QY	2858	AGAATATTGGAGACACCGTTCTTAATTCAGCTATTAAATACTCGTCTTCTTAAGCATCC	2917
Db	7606	AGAATATTGGAGACACCGTTCTTAATTCAGCTATTAAATACTCGTCTTCTTAAGCATCC	7665
QY	2918	TTCAATTCCTTTTAATAACAAATTAAGCATTAATCTTCAACAAATAGGCCGCTTGTGTA	2977
Db	7666	TTCAATTCCTTTTAATAACAAATTAAGCATTAATCTTCAACAAATAGGCCGCTTGTGTA	7725
QY	2978	ACTACTCTTTAATAAAAAATTTTTCGCTTCCCAATTCCAATTCGAATAATAGAAAATC	3037
Db	7726	ACTACTCTTTAATAAAAAATTTTTCGCTTCCCAATTCCAATTCGAATAATAGAAAATC	7785
QY	3038	CATCTTCATTCGCTTTTTCGTCATCATCTGTATGAATCAAAATCGCTTCTTCTGTGTGAT	3097
Db	7786	CATCTTCATTCGCTTTTTCGTCATCATCTGTATGAATCAAAATCGCTTCTTCTGTGTGAT	7845
QY	3098	CAAGCTTTAATTTTATGTAATTTCTTTTAAACAAACCCCATAGAGATTAACCTTTTAC	3157
Db	7846	CAAGCTTTAATTTTATGTAATTTCTTTTAAACAAACCCCATAGAGATTAACCTTTTAC	7905
QY	3158	GGTGTAAACCTTCTCCAAATCAGACAAACGTTTCCAAATCTTTTCTTCAATCGGTCAT	3217

Db	7906	GGGTAAACCTTCTCTCAAAATCAGACAAAGTTTCAAATTTCTTTTCATCATCGGTCA	7965
Qy	3218	TAAAAATCCGGTATCCTTTACAGGATATTTTGCAGTTTCGTCGAATTCGGGATTTGATATCCG	3277
Db	7966	TAAAAATCCGGTATCCTTTACAGGATATTTTGCAGTTTCGTCGAATTCGGGATTTGATATCCG	8025
Qy	3278	ATTATATTTATTTTTCGGTATTTTATTATAAACGTTCTCA	3318
Db	8026	ATTATATTTATTTTTCGGTATTTTATTATAAACGTTCTCA	8066
RESULT 12			
AAT32232			
ID	AAT32232 standard; DNA; 10140 BP.		
XX	AAT32232;		
XX	27-OCT-1996 (first entry)		
XX	Plasmid pBE92.		
XX	Thermostable enzymes; xylanase; xynA gene; beta-glucosidase; Bacillus;		
KW	Caldoecellum saccharolyticum; polymerase chain reaction; PCR; primer;		
KW	signal peptide; pBE92; ss.		
XX	Bacillus; sp.		
OS	Synthetic.		
OS	Chimeric.		
XX	WO9623987-A1.		
FN	08-AUG-1996.		
XX	24-JAN-1996; 96WO-US000891.		
PD	30-JAN-1995; 95US-00380521.		
PF	(DUPO) DU PONT DE NEMOURS & CO E I.		
XX	Jackson EN, Liu G;		
PI	WPI; 1996-371435/37.		
DR	Thermostable enzymes e.g. xylanase and beta-glucosidase - are produced		
XX	extracellularly, in absence of generally required signal peptide, and in		
PT	high yields from recombinant bacteria.		
PT	Example 1; Page 32-38; 48pp; English.		
PS	Plasmid pBE92 (AAT32232) contains the Bacillus alkaline protease gene		
XX	promoter (aprr) and signal sequence (aprss). A PCR product (see also		
CC	AAT32230-31) contg. the Caldoecellum saccharolyticum xylanase gene (xynA)		
CC	was ligated to the large fragment of pBE92. The resulting plasmid,		
CC	pBE158, contained the aprr-aprss-xynA fusion. Bacillus subtilis		
CC	transformants produced xylanase at 37 U/ml supernatant; this compared		
CC	with 272 and 110 U/ml using constructs aprr-xynA (pBE145) (see also		
CC	AAT32225-26) and nrrp-xynA (pBE145) (see also AAT32227), respectively,		
CC	which lacked a signal sequence		
XX	Sequence 10140 BP; 2821 A; 2414 C; 2129 G; 2776 T; 0 U; 0 Other;		
SQ	Query Match 29.0%; Score 1493; DB 2; Length 10140;		
	Best Local Similarity 98.1%; Pred. No. 3.2e-259;		
	Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;		
Qy	1778	TCTTATCTTGATATAAGGTAACTATTTCGGGATAGACTGTAACTTCACGCATA	1833
Db	7523	TACTCTTTTAAATCCCGACTGCGCAATGCGGATAGACTGTAACTTCACGCATA	7583
Qy	1838	AAATCCCTTTCAATTTTCTTAATGAAATCTATTACTTTATTATTAAATTCATTCGCTCAT	1899
Db	7583	AAATCCCTTTCAATTTTCTTAATGAAATCTATTACTTTATTATTAAATTCATTCGCTCAT	7643

QY 1898 AATTAAATCCTTTTCTTTATACGCAAAATGGCCGATTTAAGCACACCCCTTTATTCGGTT 1857  
Db 7643 AATTAAATCCTTTTCTTTATACGCAAAATGGCCGATTTAAGCACACCCCTTTATTCGGTT 7702  
QY 1958 AATGGCCATGACAGCCGATGTAATTAATTAATAGAGAGGTAATTAATTAATAGTAACC 2017  
Db 7703 AATGGCCATGACAGCCGATGTAATTAATTAATAGAGAGGTAATTAATTAATAGTAACC 7762  
QY 2018 AACATGATTAACAATTAATAGAGGTCATCGTTCAAAATGGTATGCGTTTTCACACATCCA 2077  
Db 7763 AACATGATTAACAATTAATAGAGGTCATCGTTCAAAATGGTATGCGTTTTCACACATCCA 7822  
QY 2078 CTATATATCCGTGTCGTCTGTCACATCTGTAATCCCATTCAGAAATTTCTAGGGATT 2137  
Db 7823 CTATATATCCGTGTCGTCTGTCACATCTGTAATCCCATTCAGAAATTTCTAGGGATT 7882  
QY 2138 CCAGAAGTTTCTCAGAGTCGGAAGTTGACAGACATTTACGAATGCGCACAGATGGTCAT 2197  
Db 7883 CCAGAAGTTTCTCAGAGTCGGAAGTTGACAGACATTTACGAATGCGCACAGATGGTCAT 7942  
QY 2198 AACCTGAAGGAGATCTGATGCTTAATCTGCTTAACGCAATTAACGCCATTTGCCCTGCATATT 2257  
Db 7943 AACCTGAAGGAGATCTGATGCTTAATCTGCTTAACGCAATTAACGCCATTTGCCCTGCATATT 8002  
QY 2258 AACAGATCGGATGATGAGACCAATCAATGCGACCTGCAATGCTACCTGTACAGTCA 2317  
Db 8003 AACAGATCGGATGATGAGACCAATCAATGCGACCTGCAATGCTACCTGTACAGTCA 8062  
QY 2318 AGGATGCTAGAAATGTTGTCGGTCTTGCACAGCAATTAACGCCATTTGCCCTGCATATT 2377  
Db 8063 AGGATGCTAGAAATGTTGTCGGTCTTGCACAGCAATTAACGCCATTTGCCCTGCATATT 8122  
QY 2378 CAACACGCTCTTACAGATAGGCAAAATCGCATCGTGGAAAGTTGGGCTTCTACCG 2437  
Db 8123 CAACACGCTCTTACAGATAGGCAAAATCGCATCGTGGAAAGTTGGGCTTCTACCG 8182  
QY 2438 ATTATGAGATTTGATACACATTTCTTAAGTATCCACCTGATCAATCAATCGGCAAAATAG 2497  
Db 8183 ATTATGAGATTTGATACACATTTCTTAAGTATCCACCTGATCAATCAATCGGCAAAATAG 8242  
QY 2498 AGAAAAATGACCATGTTAGCGGCAATCTGATTCACCTGAGATGCAATATCTAGTA 2557  
Db 8243 AGAAAAATGACCATGTTAGCGGCAATCTGATTCACCTGAGATGCAATATCTAGTA 8302  
QY 2558 GAATCTCTGCTATCAAAATTCATCTCCACCTTCCACTCACCGGTTGTCATTCATGGC 2617  
Db 8303 GAATCTCTGCTATCAAAATTCATCTCCACCTTCCACTCACCGGTTGTCATTCATGGC 8362  
QY 2618 TGAATCTGCTCTCTGTTGACATGACACATCAATCAATATCGAATAGGCGCCAT 2677  
Db 8363 TGAATCTGCTCTCTGTTGACATGACACATCAATCAATATCGAATAGGCGCCAT 8422  
QY 2678 CAGTCTGACGACCAAGAGGCAATAACCAATAGCTTAACATCATCCCATATTTAT 2737  
Db 8423 CAGTCTGACGACCAAGAGGCAATAACCAATAGCTTAACATCATCCCATATTTAT 8482  
QY 2738 CCAATATTCGTTCTTAATTTATGAAACAATCTTCAATTCCTTCTCTAGTCATTTATTA 2797  
Db 8483 CCAATATTCGTTCTTAATTTATGAAACAATCTTCAATTCCTTCTCTAGTCATTTATTA 8542  
QY 2798 TTGTCATATTCATTTCTCATCTCCCTTTTCAGATAATTTAGATTTGCTTTTCTAAATA 2857  
Db 8543 TTGTCATATTCATTTCTCATCTCCCTTTTCAGATAATTTAGATTTGCTTTTCTAAATA 8602  
QY 2858 AGAATATTTGGAGACCGGTTCTTATCTAGCTATTAAATAAATCAATCTTCTTCTAAGCATCC 2917  
Db 8603 AGAATATTTGGAGACCGGTTCTTATCTAGCTATTAAATAAATCAATCTTCTTCTAAGCATCC 8662  
QY 2918 TTCAATCTTTTAAATAACAATTAATAGATCTTAATCTTCAACAACTGGCCGCTTTGTTGA 2977  
Db 8663 TTCAATCTTTTAAATAACAATTAATAGATCTTAATCTTCAACAACTGGCCGCTTTGTTGA 8722

QY 2978 ACTACTCTTTAATAAAATAAATTTTCGGTCCCAATTCACATTCGAATTAAGAAATC 3037  
Db 8723 ACTACTCTTTAATAAAATAAATTTTCGGTCCCAATTCGAATTCGAATTAAGAAATC 8782  
QY 3038 CATCTTCATCGGCTTTTTCGTCATCATCTGATGAATCAAAATCGCCTTCTTCTGTGCAT 3097  
Db 8783 CATCTTCATCGGCTTTTTCGTCATCATCTGATGAATCAAAATCGCCTTCTTCTGTGCAT 8842  
QY 3098 CAAGGTTTAAATTTTATGTTATTTCTTTTAAACCAACCATAGGATTAACCTTTTAC 3157  
Db 8843 CAAGGTTTAAATTTTATGTTATTTCTTTTAAACCAACCATAGGATTAACCTTTTAC 8902  
QY 3158 GGTGTAAACCTTCTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGTCA 3217  
Db 8903 GGTGTAAACCTTCTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGTCA 8962  
QY 3218 TAAATATCGTATCCCTTTACAGGATTTTTCGATTTTCGATTCGATTCGATTCGATTCG 3277  
Db 8963 TAAATATCGTATCCCTTTACAGGATTTTTCGATTTTCGATTCGATTCGATTCGATTCG 9022  
QY 3278 ATTATATTTATTTTTCGGTATTTTATTTAAACGTCCTCA 3318  
Db 9023 ATTATATTTATTTTTCGGTATTTTATTTAAACGTCCTCA 9063  
RESULT 13  
AAD49392  
ID AAD49392 standard; DNA; 6661 BP.  
XX  
AC AAD49392;  
XX  
DT 07-MAR-2003 (first entry)  
XX  
DE PMOL995 plasmid DNA.  
XX  
KW Pectate lyase; EC 4.2.2.2; detergent; surfactant; cleaning; dishwashing;  
KW fabric stain removal; fabric whiteness maintenance; fabric softening;  
KW fabric colour appearance; fabric dye transfer inhibition; de-pilling;  
KW dental application; oral application; colour clarification; enzyme;  
KW clay stain removal; ds.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 1..1962  
FT /tag= a  
FT /note= "Encodes PUB110 plasmid"  
FT terminator 1963..2305  
FT /tag= b  
FT /note= "Encodes transcriptional terminator from amyL gene  
FT of B. licheniformis ATCC14580 and a few introduced  
FT restriction sites" 2306..3766  
FT misc\_feature /tag= c  
FT /note= "Encodes mature part of alpha-amyLase"  
FT complement (3767..4075)  
FT /tag= d  
FT /note= "Encodes the promoter and signalpeptide of alpha-  
FT amyLase" 4076..6661  
FT misc\_feature /tag= e  
FT /note= "Encodes PUB110 plasmid"  
XX  
FN WO200292741-A2.  
XX  
PD 21-NOV-2002.  
XX  
PF 14-MAY-2002; 2002WO-DK000315.  
XX  
PR 14-MAY-2001; 2001DK-00000755.  
XX  
PR 14-MAY-2001; 2001US-0290738P.  
XX  
PA (NOVO ) NOVOZYMES AS.

XX Eskelund MB, Schuelein M, Nielsen VS, Smets J;  
 XX WPI; 2003-059005/05.  
 XX New polypeptide useful as ingredient of detergent composition for  
 PT cleaning fabric, dishware or hard surface, encoded by DNA sequence  
 PT endogenous to strain of *Bacillus subtilis*, comprises pectate lyase  
 PT activity.  
 XX Disclosure; Page 84-87; 95pp; English.  
 XX The present invention relates to pectate lyase (EC 4.2.2.2) proteins and  
 CC polynucleotides encoding such proteins. The invention also relates to  
 CC detergent compositions comprising a surfactant and pectate lyase enzyme.  
 CC The detergent is useful for cleaning a fabric, a dishware or hard surface  
 CC to provide superior cleaning performance, for fabric cleaning and/or  
 CC fabric stain removal and/or fabric whiteness maintenance and/or fabric  
 CC softening and/or fabric colour appearance and/or fabric dye transfer  
 CC inhibition, for cleaning hard surfaces such as floors, walls or bathroom  
 CC tiles, for hand and machine dishwashing and for oral and/or dental  
 CC applications. They are useful for colour clarification, de-pilling and in  
 CC clay stain removal. The present sequence is pMOL995 plasmid DNA  
 XX Sequence 6661 BP; 1859 A; 1534 C; 991 G; 2277 T; 0 U; 0 Other;  
 SQ Query Match 29.0%; Score 1490.6; DB 7; Length 6661;  
 Best Local Similarity 99.1%; Pred. No. 8.2e-259;  
 Matches 1499; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 QY 1806 TGCGGGATAGACTGATCAATCTCCAGCAATAAATCCCTTTCATTTCTTAATGTAAT 1865  
 DB 4072 TCCCGGGATAGACTGATCAATCTCCAGCAATAAATCCCTTTCATTTCTTAATGTAAT 4131  
 QY 1866 CTATTACCTTATTATTAAATCAATTCGCTCAATAATTAATCCCTTTCATTAATGCAAAA 1925  
 DB 4132 CTATTACCTTATTATTAAATCAATTCGCTCAATAATTAATCCCTTTCATTAATGCAAAA 4191  
 QY 1926 TGCGCCGATTAAGCACACCTTTATTCGGTTAATGCGCCATGACAGCCATGATAATAC 1985  
 DB 4192 TGCGCCGATTAAGCACACCTTTATTCGGTTAATGCGCCATGACAGCCATGATAATAC 4251  
 QY 1986 TAATACTAGGAGAGCTTAATAAATACGTAACCAACATGATTAATAAATATTAGAGGTCAAT 2045  
 DB 4252 TAATACTAGGAGAGCTTAATAAATACGTAACCAACATGATTAATAAATATTAGAGGTCAAT 4311  
 QY 2046 CGTTCAAAATGGTATGCGTTTTCACATCCATATATCCGTTGCTGTCCTCACTC 2105  
 DB 4312 CGTTCAAAATGGTATGCGTTTTCACATCCATATATCCGTTGCTGTCCTCACTC 4371  
 QY 2106 CTGAATCCCATCCAGAAATCTCTAGCGATTCCAGAAATTTCTCAGAGTCGGAAGTTG 2165  
 DB 4372 CTGAATCCCATCCAGAAATCTCTAGCGATTCCAGAAATTTCTCAGAGTCGGAAGTTG 4431  
 QY 2166 ACCAGACATTTAGAACTGGCAGATGGTCAATCACTGAAGGAAGATCTGATTGCTTAAC 2225  
 DB 4432 ACCAGACATTTAGAACTGGCAGATGGTCAATCACTGAAGGAAGATCTGATTGCTTAAC 4491  
 QY 2226 TGCCTTCAAGTTAGACCGAGCCCTCGTGTATTAACAGATGCGATGATCGACCAATCAAA 2285  
 DB 4492 TGCCTTCAAGTTAGACCGAGCCCTCGTGTATTAACAGATGCGATGATCGACCAATCAAA 4551  
 QY 2286 CATGGCACTGCGATTGCTTACCTGATACAGTCAAGGATGGTAAATTTGTCGGTCCCTTG 2345  
 DB 4552 CATGGCACTGCGATTGCTTACCTGATACAGTCAAGGATGGTAAATTTGTCGGTCCCTTG 4611  
 QY 2346 CACACGAATATTACGCCATTTGCTGATATTCAACAGCTCTTCTAGATAAGGGACA 2405  
 DB 4612 CACACGAATATTACGCCATTTGCTGATATTCAACAGCTCTTCTAGATAAGGGACA 4671  
 QY 2406 AATCGCATCGTGGAGCTTTGGCTCTTACCGATTTAGCAGTTTGTATACACTTTCTCTAA 2465  
 DB 4672 AATCGCATCGTGGAGCTTTGGCTCTTACCGATTTAGCAGTTTGTATACACTTTCTCTAA 4731

QY 2466 GTATCCACCTGAATCATATAATCGCAAAATAGAAAAATAGAACATGTGTAAGCGGCA 2525  
 DB 4732 GTATCCACCTGAATCATATAATCGCAAAATAGAAAAATAGAACATGTGTAAGCGGCA 4791  
 QY 2526 ATCTGATTCACCTGAGATGATATCTAGTAGAATCTCTCGCTATCAAAATTCACATTC 2585  
 DB 4792 ATCTGATTCACCTGAGATGATATCTAGTAGAATCTCTCGCTATCAAAATTCACATTC 4851  
 QY 2586 CACCTTCCACTCACCGGTTGCTCATTTCATGGCTGAACCTCTGCTCTCTGTTGACATGAC 2645  
 DB 4852 CACCTTCCACTCACCGGTTGCTCATTTCATGGCTGAACCTCTGCTCTCTGTTGACATGAC 4911  
 QY 2646 ACACATCATCTCAATATCCGAATAGGGCCCATCTCTGACGACCAAGAGCCATAAAC 2705  
 DB 4912 ACACATCATCTCAATATCCGAATAGGGCCCATCTCTGACGACCAAGAGCCATAAAC 4971  
 QY 2706 ACCAATAGCTTTAAACATCATATCCCATATTTATCCCAATATTCCTTAAATTCATGAC 2765  
 DB 4972 ACCAATAGCTTTAAACATCATATCCCATATTTATCCCAATATTCCTTAAATTCATGAC 5031  
 QY 2766 AATCTTCATCTCTCTCTAGTCAATATTTATTTGGTCAATTCACATCTCTATCCCTT 2825  
 DB 5032 AATCTTCATCTCTCTCTAGTCAATATTTATTTGGTCAATTCACATCTCTATCCCTT 5091  
 QY 2826 TTCAGATATTTTAGATTTGCTTTTCTAATAAGAAATATTTGAGAGCACCTGTTCTTATT 2885  
 DB 5092 TTCAGATATTTTAGATTTGCTTTTCTAATAAGAAATATTTGAGAGCACCTGTTCTTATT 5151  
 QY 2886 CAGCTATTAAATACCTGCTCTCTAAGCATCTCTCAATCCCTTTAATAACAATATAGCA 2945  
 DB 5152 CAGCTATTAAATACCTGCTCTCTAAGCATCTCTCAATCCCTTTAATAACAATATAGCA 5211  
 QY 2946 TCTAATCTTCAACAACCTGCGGTTGTTGAACTACTCTTTAATAAATAAATATTTTCG 3005  
 DB 5212 TCTAATCTTCAACAACCTGCGGTTGTTGAACTACTCTTTAATAAATAAATATTTTCG 5271  
 QY 3006 TTCCCAATTTCCCAATTTGCAATTAAGAAATCCATCTTCATCGGCTTTTTCGTCATCATC 3065  
 DB 5272 TTCCCAATTTCCCAATTTGCAATTAAGAAATCCATCTTCATCGGCTTTTTCGTCATCATC 5331  
 QY 3066 TGTATGAATCAATCGCTCTCTCTGTCATCAAGGTTAATTTTATGATTTCTTT 3125  
 DB 5332 TGTATGAATCAATCGCTCTCTCTGTCATCAAGGTTAATTTTATGATTTCTTT 5391  
 QY 3126 TAACAACACCATAGAGATTAACCTTTACGGTGTAAACCTTCCTCCAAATCAGACAA 3185  
 DB 5392 TAACAACACCATAGAGATTAACCTTTACGGTGTAAACCTTCCTCCAAATCAGACAA 5451  
 QY 3186 ACCTTTCAAAATCTTTTCTTCATCATCGGTCAATAAATCCGTATCCTTTACAGGATATT 3245  
 DB 5452 ACCTTTCAAAATCTTTTCTTCATCATCGGTCAATAAATCCGTATCCTTTACAGGATATT 5511  
 QY 3246 TGCAGTTTCGTCAATTCGGGATTTGATATCCGATTTATATTTTTCGGTATTTTAA 3305  
 DB 5512 TGCAGTTTCGTCAATTCGGGATTTGATATCCGATTTATATTTTATTTTTCGGTATTTTAA 5571  
 QY 3306 TTAACACGCTCA 3318  
 DB 5572 TTGAACCTTTTACA 5584  
 RESULT 14  
 AAA98012  
 ID AAA98012 standard; DNA; 5302 BP.  
 XX  
 AC AAA98012;  
 XX  
 DT 05-FEB-2001 (first entry)  
 XX Expression vector p602/-CAT DNA.  
 XX Lumazine synthase; capsid; cytostatic; antiviral; antibacterial; vaccine;



gene therapy; immunotherapy; biosensor; diagnosis; ds.

Synthetic.

WO200053229-A2.

14-SEP-2000.

03-MAR-2000; 2000WO-EP001899.

08-MAR-1999; 99DE-01010102.

(FISC/) FISCHER M.

(BACH/) BACHER A.

Fischer M, Bacher A;

WPI; 2000-572230/53.

Protein conjugate based on lumazine synthase as carrier, useful e.g. for vaccination and immunotherapy, contains many functional molecules attached to outside of the carrier.

Example 1; Page 130-132; 180pp; German.

This invention describes a novel protein conjugate (I) comprising at least 1 functional region (FR) at any position in the sequence of a carrier protein (II) to form a capsid three-dimensional structure of a lumazine synthase (LS) type, such that the outer periphery is covalently linked to many FR. The invention also describes (1) a hetero-oligomeric protein conjugate (Ia) comprising either a mixture of at least 2 different (I) or of at least 1 (I) and at least 1 (II) that lacks FR, with the components optionally covalently linked by chemical treatment; (2) a method for preparation of (I) and (Ia); (3) a vector for preparation of (I); (4) a DNA (III) that encodes (I); (5) LS from *Bacillus subtilis* with Cys 93 and/or Cys139 replaced by Ser; (6) DNA (IIia) encoding LS from *Aquifex aeolicus* which is codon-optimized for expression in a recombinant *Escherichia coli* strain; (7) a chimeric protein (CP) comprising amino acids (aa) 1-60 from LS of *B. subtilis* and aa 61-154 from LS of *A. aeolicus* for use as (II); and (8) pharmaceutical compositions and vaccines containing (I) and (Ia). The products of the invention have cytostatic, antiviral and antibacterial activity and can be used in a vaccine or for gene therapy. (I) are used as pharmaceuticals (e.g. for immunotherapy of human immune deficiency virus infection or tumors) and in vaccines, including multivalent vaccines, against bacterial or viral infections, to produce diagnostic or therapeutic antibodies, for selective detection, purification and characterization of antibodies, and for preparation of protein libraries. (I) may also be used as biosensors and for diagnosis, e.g. of tumors. Genes that encode (I) are useful in DNA vaccines and for preparation of plant-based oral vaccines. (I) can contain many FR (same or different) at the surface of a spherical particle (LS comprises 60 subunits that assemble into an icosahedron). The large number of FR may increase sensitivity in immunoassays and the efficiency of immunotherapy agents

Sequence 5302 BP; 1502 A; 1208 C; 962 G; 1630 T; 0 U; 0 Other;

Query Match 28.9%; Score 1487.6; DB 3; Length 5302;

Best Local Similarity 99.1%; Pred. No. 2.8e-258;

Matches 1496; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY	1809	CGGGATAGACTGTAACTCTCAGCATATAAATCCCTTTTCATTTCTAATGTAATCTA	1868
DB	2716	CGGGATAGACTGTAACTCTCAGCATATAAATCCCTTTTCATTTCTAATGTAATCTA	2775
QY	1869	TTACCTTTAATTAATTCATTCGCTCATATTAATCTTTTCTTATTACGCAAAATGG	1928
DB	2776	TTACCTTTAATTAATTCATTCGCTCATATTAATCTTTTCTTATTACGCAAAATGG	2835
QY	1929	CCCGATTTAAGACACACCTTTATTCCGTTAATGGCCATGACGCCATGATTAATCTAA	1988
DB	2836	CCCGATTTAAGACACACCTTTATTCCGTTAATGGCCATGACGCCATGATTAATCTAA	2895
QY	1989	TACTAGGAGAGTTAATAAATACGTAAACCAATGATTAAACAATTTATAGAGTCACTGT	2048
DB	2896	TACTAGGAGAGTTAATAAATACGTAAACCAATGATTAAACAATTTATAGAGTCACTGT	2955
QY	2049	TCAAAATGGTATCGCTTTTGACACATCCACTATATATCCGTGCTGCTGCTGCTGCTG	2108
DB	2956	TCAAAATGGTATCGCTTTTGACACATCCACTATATATCCGTGCTGCTGCTGCTGCTG	3015
QY	2109	AATCCCATTTCCAGAAATTTCTTAGCGATTCCAGAGTTTCTCAGAGTCGGAAGTTGACC	2168
DB	3016	AATCCCATTTCCAGAAATTTCTTAGCGATTCCAGAGTTTCTCAGAGTCGGAAGTTGACC	3075
QY	2169	AGACATTACGAATCTGGCAGAGATGCTCATACCTGAAGAGATCTGATTCTTAATCTGC	2228
DB	3076	AGACATTACGAATCTGGCAGAGATGCTCATACCTGAAGAGATCTGATTCTTAATCTGC	3135
QY	2229	TTTCAGTTAAGACCGGAGCGCTCGTGTATTAACAGATGCGATGCGACCAATCAACAT	2288
DB	3136	TTTCAGTTAAGACCGGAGCGCTCGTGTATTAACAGATGCGATGCGACCAATCAACAT	3195
QY	2289	GGCAGCTGCAATGCTACTGTACAGTCAAGATGCTAGAAATGTTGCGTCTCTTGAC	2348
DB	3196	GGCAGCTGCAATGCTACTGTACAGTCAAGATGCTAGAAATGTTGCGTCTCTTGAC	3255
QY	2349	ACGAATATTACGCCATTTGCGCTGCATATTCAAAACAGCTCTTCTACGATAAGGCAAAAT	2408
DB	3256	ACGAATATTACGCCATTTGCGCTGCATATTCAAAACAGCTCTTCTACGATAAGGCAAAAT	3315
QY	2409	CCGATCGTGAACGTTTGGGCTTTACCGATTTAGCAGTTTGATACACTTTCTCTAAGTA	2468
DB	3316	CCGATCGTGAACGTTTGGGCTTTACCGATTTAGCAGTTTGATACACTTTCTCTAAGTA	3375
QY	2469	TCGACTGCAATCAATAATCGCCAAATAGAGAAATTTGACCATGTGTAAAGCGCCAATC	2528
DB	3376	TCGACTGCAATCAATAATCGCCAAATAGAGAAATTTGACCATGTGTAAAGCGCCAATC	3435
QY	2529	TGATTCCACCTGAGATGATTAATCTAGTAGAATCTCTTCGCTATCAAAATTCACCTCCAC	2588
DB	3436	TGATTCCACCTGAGATGATTAATCTAGTAGAATCTCTTCGCTATCAAAATTCACCTCCAC	3495
QY	2589	CTTCACCTACCGGTTGCCATTCATGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2648
DB	3496	CTTCACCTACCGGTTGCCATTCATGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3555
QY	2649	CATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATTAACACC	2708
DB	3556	CATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATTAACACC	3615
QY	2709	AATAGCCTTAACATCATCCCATATTTATCAATATTCGTTCCCTTAATTTATGAGCAAT	2768
DB	3616	AATAGCCTTAACATCATCCCATATTTATCAATATTCGTTCCCTTAATTTATGAGCAAT	3675
QY	2769	CTTCATTTCTTTCTCTTAGTCATTTATTATTTGGTCCATTCATCTCTCTCTCTCTCTCTC	2828
DB	3676	CTTCATTTCTTTCTCTTAGTCATTTATTATTTGGTCCATTCATCTCTCTCTCTCTCTC	3735
QY	2829	AGATAATTTTAGATTTGCTTTTCTAAATAAGAAATTTTGGAGAGACCGGTTCTTATTTCAG	2888
DB	3736	AGATAATTTTAGATTTGCTTTTCTAAATAAGAAATTTTGGAGAGACCGGTTCTTATTTCAG	3795
QY	2889	CTATTAAATACCTGCTTCTCCCTAAGCATCTTCAATCCCTTTTAAATAAATAATATAGCATCT	2948
DB	3796	CTATTAAATACCTGCTTCTCCCTAAGCATCTTCAATCCCTTTTAAATAAATAATATAGCATCT	3855
QY	2949	AATCTTCAACAACTGGCCGCTTTTGTGCACTACTCTTTTAAATAAATAATTTTCCGTTTC	3008
DB	3856	AATCTTCAACAACTGGCCGCTTTTGTGCACTACTCTTTTAAATAAATAATTTTCCGTTTC	3915
QY	3009	CCAAATCCACATGCAATAATAGAAATCCCATCTTCATCGGCTTTTTCGTCATCATCTGT	3068
DB	3916	CCAAATCCACATGCAATAATAGAAATCCCATCTTCATCGGCTTTTTCGTCATCATCTGT	3975
QY	3069	ATGAATCAAAATCGCCTTCTCTGTGTCATCAAGGTTTAAATTTTATGTATTTCTTTTAA	3128

Db 3976 ATGAATCAATCGGCTTCTCTGTGTCATCAAGGTTAAATTTTATGTATTTCTTTAA 4035  
 Qy 3129 CAACCCACCATAGAGATTAACTTTACGGTGTAAACCTTCCCAAAATCAGACAAACG 3188  
 Db 4036 CAACCCACCATAGAGATTAACTTTACGGTGTAAACCTTCCCAAAATCAGACAAACG 4095  
 Qy 3189 TTTCAAATCTTTCTTTCATCATCGGTGCATATAAATCCGATATCTTTACAGATATTTGC 3248  
 Db 4096 TTTCAAATCTTTCTTTCATCATCGGTGCATATAAATCCGATATCTTTACAGATATTTGC 4155  
 Qy 3249 AGTTTCGTCATTCGGATTCGTATATCGGATTTATTTATTTTCGTTATTTTATTA 3308  
 Db 4156 AGTTTCGTCATTCGGATTCGTATATCGGATTTATTTATTTTCGTTTCGATCAATTTG 4215  
 Qy 3309 AAACGTCCTCA 3318  
 Db 4216 AACTTTTACA 4225  
 RESULT 15  
 AAA98014  
 ID AAA98014 standard; DNA; 5767 BP.  
 AC AAA98014;  
 XX 05-FEB-2001 (first entry)  
 DT Expression vector p602-BS-LuSy DNA.  
 DE Lumazine synthase; capsid; cytostatic; antiviral; antibacterial; vaccine;  
 KW gene therapy; immunotherapy; biosensor; diagnosis; ds.  
 KW Synthetic.  
 OS  
 XX WO200053229-A2.  
 XX 14-SEP-2000.  
 XX 03-MAR-2000; 2000WO-EP001899.  
 XX 08-MAR-1999; 99DE-01010102.  
 XX (FISC/) FISCHER M.  
 XX (BACH/) BACHER A.  
 XX Fischer M, Bacher A;  
 XX WPI; 2000-572230/53.  
 XX Protein conjugate based on lumazine synthase as carrier, useful e.g. for  
 PT vaccination and immunotherapy, contains many functional molecules  
 PT attached to outside of the carrier.  
 XX Example 2; Page 134-137; 180pp; German.  
 XX This invention describes a novel protein conjugate (I) comprising at  
 CC least 1 functional region (FR) at any position in the sequence of a  
 CC carrier protein (II) to form a capsid three-dimensional structure of the  
 CC lumazine synthase (LS) type, such that the outer periphery is covalently  
 CC linked to many FR. The invention also describes (1) a hetero-oligomeric  
 CC protein conjugate (Ia) comprising either a mixture of at least 2  
 CC different (i) or of at least 1 (ii) and at least 1 (iii) that lacks FR,  
 CC with the components optionally covalently linked by chemical treatment;  
 CC (2) a method for preparation of (I) and (Ia); (3) a vector for  
 CC preparation of (I); (4) a DNA (III) that encodes (I); (5) LS from  
 CC Bacillus subtilis with Cys 93 and/or Cys139 replaced by Ser; (6) DNA  
 CC (IIa) encoding LS from Aquifex aeolicus which is codon-optimized for  
 CC expression in a recombinant Escherichia coli strain; (7) a chimeric  
 CC protein (CP) comprising amino acids (aa) 1-60 from LS of B. subtilis and  
 CC aa 61-154 from LS of A. aeolicus for use as (II); and (8) pharmaceutical  
 CC compositions and vaccines containing (I) and (Ia). The products of the  
 CC invention have cytostatic, antiviral and antibacterial activity and can

CC be used in a vaccine or for gene therapy. (I) are used as pharmaceuticals  
 CC (e.g. for immunotherapy of human immune deficiency virus infection or  
 CC tumors), including multivalent vaccines, against  
 CC bacterial or viral infections, to produce diagnostic or therapeutic  
 CC antibodies, for selective detection, purification and characterization of  
 CC antibodies, and for preparation of protein libraries. (I) may also be  
 CC used as biosensors and for diagnosis, e.g. of tumors. Genes that encode  
 CC (I) are useful in DNA vaccines and for preparation of plant-based oral  
 CC vaccines. (I) can contain many FR (same or different) at the surface of a  
 CC spherical particle (US comprises 60 subunits that assemble into an  
 CC icosahedron). The large number of FR may increase sensitivity in  
 CC immunoassays and the efficiency of immunotherapy agents  
 XX  
 XX Sequence 5767 BP; 1654 A; 1297 C; 1073 G; 1743 T; 0 U; 0 Other;  
 Qy Query Match 28.9%; Score 1487.6; DB 3; Length 5767;  
 Db Best Local Similarity 99.1%; Pred. No. 2.8e-258;  
 Mismatches 1496; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 Qy 1809 CGGATAGACTGTAAACATTTCTCAGCATATAAATCCCTTTTCATTTTCTTAATGTAATCTA 1868  
 Db 3181 CGGATAGACTGTAAACATTTCTCAGCATATAAATCCCTTTTCATTTTCTTAATGTAATCTA 3240  
 Qy 1869 TTACCTTATTATTAAATTCGCTCAATAATTAATCCTTTTCTTTTATTACGAAATGG 1928  
 Db 3241 TTACCTTATTATTAAATTCGCTCAATAATTAATCCTTTTCTTTTATTACGAAATGG 3300  
 Qy 1929 CCGATTTAAGCACACCTTTATTCGGTTAATGCGCATGAGAGCCATGATTAATCTAA 1988  
 Db 3301 CCGATTTAAGCACACCTTTATTCGGTTAATGCGCATGAGAGCCATGATTAATCTAA 3360  
 Qy 1989 TACTAGGAGAGTTAATAAATACGTAACCAACATGATTAACCAATTAATAGAGTCAATCGT 2048  
 Db 3361 TACTAGGAGAGTTAATAAATACGTAACCAACATGATTAACCAATTAATAGAGTCAATCGT 3420  
 Qy 2049 TCAAAATGTTATCGTTTTCACACATCCACTATATATCGTGTCTCTGTCCACTCCCTG 2108  
 Db 3421 TCAAAATGTTATCGTTTTCACACATCCACTATATATCGTGTCTCTGTCCACTCCCTG 3480  
 Qy 2109 AATCCCATCCAGAAATTTCTTAGCGATTCCAGAGTTTCTCAGAGTCGGAAGTTGACC 2168  
 Db 3481 AATCCCATCCAGAAATTTCTTAGCGATTCCAGAGTTTCTCAGAGTCGGAAGTTGACC 3540  
 Qy 2169 AGACATTACGAATGGCACAGATGGTCAATACTGAAGAGATCTGTATTGCTTAATGTC 2228  
 Db 3541 AGACATTACGAATGGCACAGATGGTCAATACTGAAGAGATCTGTATTGCTTAATGTC 3600  
 Qy 2229 TTCAGTTAAGACCGAAGCGCTCGTGTATAACAGATGCGATGATGAGACCAATCAACAT 2288  
 Db 3601 TTCAGTTAAGACCGAAGCGCTCGTGTATAACAGATGCGATGATGAGACCAATCAACAT 3660  
 Qy 2289 GGCACCTGCCATTGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTGGTCTTGCAC 2348  
 Db 3661 GGCACCTGCCATTGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTGGTCTTGCAC 3720  
 Qy 2349 ACGAATATTACGCCATTTGCTGCTCATATTCAACACAGCTTCTTACGATAGGCGCAAAAT 2408  
 Db 3721 ACGAATATTACGCCATTTGCTGCTCATATTCAACACAGCTTCTTACGATAGGCGCAAAAT 3780  
 Qy 2409 CGCATCGTGAACGGTTTGGGCTTCTACCGATTTTAGCAGTTTGTATACACATTTCTTAAGTA 2468  
 Db 3781 CGCATCGTGAACGGTTTGGGCTTCTACCGATTTTAGCAGTTTGTATACACATTTCTTAAGTA 3840  
 Qy 2469 TCCACCTGAATCAATAATCGGCAAAATAGAGAAAATTTGACCAATGTTAGCGGCCCAATC 2528  
 Db 3841 TCCACCTGAATCAATAATCGGCAAAATAGAGAAAATTTGACCAATGTTAGCGGCCCAATC 3900  
 Qy 2529 TGATTTCCACCTGAGATGATTAATCTAGTAAATCTCTTCGCTATCAAAATTTCACTTCCAC 2588  
 Db 3901 TGATTTCCACCTGAGATGATTAATCTAGTAAATCTCTTCGCTATCAAAATTTCACTTCCAC 3960  
 Qy 2589 CTTCACCTACCGGTTGTCCATTCATGGCTGAATCTGTCTTCCTTCTGTTGACATGACACA 2648



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 23, 2004, 11:31:21 ; Search time 12724.2 Seconds  
(without alignments)  
17515.470 Million cell updates/sec

Title: US-10-030-390-1  
Perfect score: 5142  
Sequence: 1 gaattcgagctggtaccg.....ccgtctgtgcttcccttaag 5142

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_atg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_on.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5142	100.0	5142	6	AX069289 Sequence
C 2	2467.2	48.0	8119	6	AR002333 Sequence
C 3	2467.2	48.0	8119	6	AR018032 Sequence
C 4	2467.2	48.0	8119	6	AR055898 Sequence
C 5	1650.4	32.1	2059	6	A09338 S.lactis pl
C 6	1650.4	32.1	2059	6	A09339 S.lactis pl
C 7	1641.6	31.9	8082	12	AF076212 Promoter
C 8	1641.6	31.9	8142	12	AF076213 Promoter
C 9	1637	31.8	2178	1	LR389A L.lactis pl
C 10	1635.2	31.8	5126	12	AF372620 GlnQ allele
C 11	1635.2	31.8	9021	12	AF061336 Tn317 del
C 12	1635.2	31.8	13561	12	AY028776 Tnpho2 mu
C 13	1629.6	31.7	2505	1	LR389A L.lactis pl
C 14	1575.2	30.6	4438	12	AB055651 Thermosen
C 15	1575.2	30.6	4438	12	AB055652 Thermosen
C 16	1575.2	30.6	4506	12	AB055650 Thermosen
C 17	1572.2	30.6	3522	12	AB084785 Transposa
C 18	1572.2	30.6	3618	12	AB084784 Transposa
C 19	1493	29.0	4548	1	PHOKANRCG
C 20	1493	29.0	7336	6	AB08030
C 21	1493	29.0	7336	6	AB08030
C 22	1493	29.0	7336	6	AB08030
C 23	1493	29.0	7336	6	AB08030
C 24	1493	29.0	7336	6	AB08030
C 25	1493	29.0	7336	6	AB08030
C 26	1493	29.0	7336	6	AB08030
C 27	1493	29.0	7336	6	AB08030
C 28	1491.4	29.0	4525	1	BSREPB
C 29	1491.4	29.0	4525	1	BSREPB
C 30	1491.4	29.0	7326	6	AB364483
C 31	1491.4	29.0	11887	1	PTITERIMOB
C 32	1491	29.0	7336	6	AB05160
C 33	1490.6	29.0	6661	6	AX601433
C 34	1490.2	29.0	4780	12	ACVPBD64
C 35	1487.6	28.9	5767	6	AX035965
C 36	1482.6	28.8	2763	6	E05086
C 37	1482.6	28.8	2763	6	E05087
C 38	1479	28.8	5609	6	AR369757
C 39	1479	28.8	5609	6	AR369758
C 40	1479	28.8	7026	6	AR369761
C 41	1479	28.8	7344	6	AR369759
C 42	1436.2	27.9	1600	1	LPCOPAREP
C 43	1386.6	27.0	5126	12	AF372620
C 44	1336.6	26.0	2266	1	SA110KAR
C 45	1335.6	26.0	2262	1	BSPRBHK

ALIGNMENTS

RESULT 1  
AX069289

LOCUS  
DEFINITION  
ACCESSION  
VERSION

KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1  
Hans.W.C., Steidler,L. and Remaut,E.R.  
Delivery of trefoil peptides  
Patent: WO 0102570-A 111-JAN-2001:  
Vlaams Interuniversitair Instituut voor Biotechnologie (BE)

AX069289 Sequence 1 from Patent WO0102570.  
AX069289,1 GI:12579162  
5142 bp DNA linear PAT 25-JAN-2001

Pred. No. is the number of results predicted by chance to have a

FEATURES		Location/Qualifiers	
source	1..5142		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
/note="Synthetic construct : plasmid"			
ORIGIN			
Query Match	100.0%;	Score 5142;	DB 6; Length 5142;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 5142;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	GAATTCAGCTCGGTACCGGGGATCTCGATCCCGGAAATTAATACGACTCACTATAGG	60
Db	1	GAATTCAGCTCGGTACCGGGGATCTCGATCCCGGAAATTAATACGACTCACTATAGG	60
Qy	61	GAGACCACAAACGGTTCCCTCTAGAAATAATTTTGGTTTAACTTTAAGAGGAGATATACA	120
Db	61	GAGACCACAAACGGTTCCCTCTAGAAATAATTTTGGTTTAACTTTAAGAGGAGATATACA	120
Qy	121	TATGAAAAAAGATATCTCAGCTATTTTAAATGCTCTACAGTCATATCTTCTGCTCGACG	180
Db	121	TATGAAAAAAGATATCTCAGCTATTTTAAATGCTCTACAGTCATATCTTCTGCTCGACG	180
Qy	181	CCCGTTGTCAGGTGTTTACGCCAGGCCAGGCCAGGCCAGGCCAGGAAACAATGTATCAT	240
Db	181	CCCGTTGTCAGGTGTTTACGCCAGGCCAGGCCAGGCCAGGAAACAATGTATCAT	240
Qy	241	GGCCCCCGGAGAGGATAAATGTGGCTTCCCGGTGTACCGCCAGCAGTGCACGGA	300
Db	241	GGCCCCCGGAGAGGATAAATGTGGCTTCCCGGTGTACCGCCAGCAGTGCACGGA	300
Qy	301	GAGAGGTGCTGTTTTGATGACAGGTGCGGGGATTCGGGTGTTCCACCCCATGGC	360
Db	301	GAGAGGTGCTGTTTTGATGACAGGTGCGGGGATTCGGGTGTTCCACCCCATGGC	360
Qy	361	CATCGAGAACACTCAAGAGAGAAATGTCCTTTCTAACTAGTAGATCCGGCTGCTAACAA	420
Db	361	CATCGAGAACACTCAAGAGAGAAATGTCCTTTCTAACTAGTAGATCCGGCTGCTAACAA	420
Qy	421	AGCCGGAAGGAAGCTGAGTTGGCTGTCGCACCGCTGAGCATAACTAGCATTAACCCCT	480
Db	421	AGCCGGAAGGAAGCTGAGTTGGCTGTCGCACCGCTGAGCATAACTAGCATTAACCCCT	480
Qy	481	TGGGGCTCTAAACGGGTCTTGGGGTTTTTTCTGCAAGGAGGAACATATAATCCGGATG	540
Db	481	TGGGGCTCTAAACGGGTCTTGGGGTTTTTTCTGCAAGGAGGAACATATAATCCGGATG	540
Qy	541	ACCTGAGGATGCAAGCTTGGCACTGGCGTCTGTTTTCACAGCTGCTGACTGGGAAAC	600
Db	541	ACCTGAGGATGCAAGCTTGGCACTGGCGTCTGTTTTCACAGCTGCTGACTGGGAAAC	600
Qy	601	CCTGGCGTTTACCCAACTTAATCGCCTTGCAGCACAATCCCTTTTCGCCAGCTGATTTTCA	660
Db	601	CCTGGCGTTTACCCAACTTAATCGCCTTGCAGCACAATCCCTTTTCGCCAGCTGATTTTCA	660
Qy	661	TTTTTGCATCTTACAACTGCAATCATATGTAATGCTCCTTTTATGTTAGTGGCACA	720
Db	661	TTTTTGCATCTTACAACTGCAATCATATGTAATGCTCCTTTTATGTTAGTGGCACA	720
Qy	721	ATGTGAGGATTTTTCGCTTTCGGCGAGGCTAGTTTACCTTTAAGTTATTTGTTATGACT	780
Db	721	ATGTGAGGATTTTTCGCTTTCGGCGAGGCTAGTTTACCTTTAAGTTATTTGTTATGACT	780
Qy	781	GGTTTAAAGCGCAAAAAAGTGTCTTTTCGTAACCTATTAATGTATCGTTTATGAAAAAC	840
Db	781	GGTTTAAAGCGCAAAAAAGTGTCTTTTCGTAACCTATTAATGTATCGTTTATGAAAAAC	840
Qy	841	GACTGTAAAGATACAGTGGCATTTATCTCATATTAAGAGCCAGTCAATTAGGCCTTATC	900
Db	841	GACTGTAAAGATACAGTGGCATTTATCTCATATTAAGAGCCAGTCAATTAGGCCTTATC	900
Qy	901	TGACAAATTCCTGAATAGAGTTTATATAACAATCTCTGCATGATAACCATCAAAACGAAATG	960

Db	901	TGACAAATTCCTGATAGATTATATAAATACTCTGCATGATACCATCAAAACAGATG	960
Qy	961	ATGTACCTGTAAAGATAGCGTAAATATATTGAAATTAACCTTTAATGAAATTTCTCTG	1020
Db	961	ATGTACCTGTAAAGATAGCGTAAATATATTGAAATTAACCTTTAATGAAATTTCTCTG	1020
Qy	1021	TGTAATATGGGTAGAGGTAATTTACTATTATTATTTGATATTTAAGTTAAACCCAGTAA	1080
Db	1021	TGTAATATGGGTAGAGGTAATTTACTATTATTATTTGATATTTAAGTTAAACCCAGTAA	1080
Qy	1081	TGAAGTCCATGGAATAATAGAAAGAGAAAGCAATTTTCAGGTATAGTGTGTTGGGAAA	1140
Db	1081	TGAAGTCCATGGAATAATAGAAAGAGAAAGCAATTTTCAGGTATAGTGTGTTGGGAAA	1140
Qy	1141	CAATTTCCCGAACCATTATATTTCTCTACATCAGAAAGGTATAAATCATAAACCTTTT	1200
Db	1141	CAATTTCCCGAACCATTATATTTCTCTACATCAGAAAGGTATAAATCATAAACCTTTT	1200
Qy	1201	GAAGTCAATCTTTACAGAGTCCAAATACAGAGAAATGTTTATAGATACACCATCAAAAT	1260
Db	1201	GAAGTCAATCTTTACAGAGTCCAAATACAGAGAAATGTTTATAGATACACCATCAAAAT	1260
Qy	1261	TGTATAAGTGGCTCTAACTTATCCCAATAACCTTAACCTCCGTCGCTATTGTAACCA	1320
Db	1261	TGTATAAGTGGCTCTAACTTATCCCAATAACCTTAACCTCCGTCGCTATTGTAACCA	1320
Qy	1321	TCTAAAGCTGTATTTGAGTTTATCGGTATAAACAATAAATCAATTTCTGTGTTAT	1380
Db	1321	TCTAAAGCTGTATTTGAGTTTATCGGTATAAACAATAAATCAATTTCTGTGTTAT	1380
Qy	1381	TTTATATCTCTGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTT	1440
Db	1381	TTTATATCTCTGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTT	1440
Qy	1441	ACTAAAGTCTGTTTGTGTTCAATTAATGATTAATATCTCTTCTCTCTCAATGTC	1500
Db	1441	ACTAAAGTCTGTTTGTGTTCAATTAATGATTAATATCTCTTCTCTCTCAATGTC	1500
Qy	1501	TAAATCAATTTTATAAAGTTTCAATTCATATGATGCTCTCTAAATTTTATCTAAAGTGA	1560
Db	1501	TAAATCAATTTTATAAAGTTTCAATTCATATGATGCTCTCTAAATTTTATCTAAAGTGA	1560
Qy	1561	TAGAGGCTTACTGCTGCTTTCTTCAATTAATGATGATGCTCTCTAAATTTTATCTAAAGTGA	1620
Db	1561	TAGAGGCTTACTGCTGCTTTCTTCAATTAATGATGATGCTCTCTAAATTTTATCTAAAGTGA	1620
Qy	1621	ACTGTAAATATAATATATTTTAAATAATCCCACTTATCCCAATTTTCTGTTGTTGAA	1680
Db	1621	ACTGTAAATATAATATATTTTAAATAATCCCACTTATCCCAATTTTCTGTTGTTGAA	1680
Qy	1681	CTAATGGGTGCTTTAGTTGAAGATAAAGACCAATTTAAATAATGTCGCTTTTGTGTTT	1740
Db	1681	CTAATGGGTGCTTTAGTTGAAGATAAAGACCAATTTAAATAATGTCGCTTTTGTGTTT	1740
Qy	1741	TTTTTAAAGGATTTGAGCGTAGCGGAAATCTCTTTTCTTTCTTATCTTTGATAAAGGTA	1800
Db	1741	TTTTTAAAGGATTTGAGCGTAGCGGAAATCTCTTTTCTTTCTTATCTTTGATAAAGGTA	1800
Qy	1801	ACTATTGCGGGATAGACTGTAACTTCTCAGCATAAATCCCTTTCAATTTCTTAATG	1860
Db	1801	ACTATTGCGGGATAGACTGTAACTTCTCAGCATAAATCCCTTTCAATTTCTTAATG	1860
Qy	1861	TAAATCTATTCCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1920
Db	1861	TAAATCTATTCCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1920
Qy	1921	CAAAATGCGCGGATTTAAGCACACCTTTTATTCGTTTAAATGCGGCATGACAGCATGATA	1980
Db	1921	CAAAATGCGCGGATTTAAGCACACCTTTTATTCGTTTAAATGCGGCATGACAGCATGATA	1980
Qy	1981	ATTACTAATCTAGAGAGTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2040

Db 1981 ATTACTAATACTAGGAGAGTTAATAAATACGTAAACCAACATGATTAAACAATTTAGAG 2040  
Qy  
2041 GTCATCGTTCAAATAGTATGCGTTTGTGACACATCCACTATATATCCGTGTCGTCTGTC 2100  
Db  
2041 GTCATCGTTCAAATAGTATGCGTTTGTGACACATCCACTATATATCCGTGTCGTCTGTC 2100  
Qy  
2101 CACTCTGTAATCCGATCCGAGAAATTCCTAGCGATCCGAGATTCCTCAGAGTCGGA 2160  
Db  
2101 CACTCTGTAATCCGATCCGAGAAATTCCTAGCGATCCGAGATTCCTCAGAGTCGGA 2160  
Qy  
2161 AGTTGACGACATTTACGAATGCGACAGATGTCATAACCTGAAGGAAGATCTGATTGC 2220  
Db  
2161 AGTTGACGACATTTACGAATGCGACAGATGTCATAACCTGAAGGAAGATCTGATTGC 2220  
Qy  
2221 TTAATCTGCTTCAAGTAAAGACGAGCGCTCGTGTATATAACAGATGCGATCCAGACCA 2280  
Db  
2221 TTAATCTGCTTCAAGTAAAGACGAGCGCTCGTGTATATAACAGATGCGATCCAGACCA 2280  
Qy  
2281 ATCAACATGCGACCTGCTGCTTACCTGTACAGTCAAGGATGTAAGATGTTGCGGT 2340  
Db  
2281 ATCAACATGCGACCTGCTGCTTACCTGTACAGTCAAGGATGTAAGATGTTGCGGT 2340  
Qy  
2341 CTTTGCACAGAAATTTACGCCATTTCCCTGCGATTTCAAACAGCTCTTCTAGCATAAGG 2400  
Db  
2341 CTTTGCACAGAAATTTACGCCATTTCCCTGCGATTTCAAACAGCTCTTCTAGCATAAGG 2400  
Qy  
2401 GCACAAATCGCATCGTGGAAAGTTGGCTTCTACCGATTTAGCAGTTTGATACACTTC 2460  
Db  
2401 GCACAAATCGCATCGTGGAAAGTTGGCTTCTACCGATTTAGCAGTTTGATACACTTC 2460  
Qy  
2461 TCTAAGTATCCACTGGAATCATAAATCGGCAAAATAGAGAAATGACCATGTGAAGC 2520  
Db  
2461 TCTAAGTATCCACTGGAATCATAAATCGGCAAAATAGAGAAATGACCATGTGAAGC 2520  
Qy  
2521 GGCCAAATCTGATCCACTGAGATGCAATCTAGTAGAATCTCTCGCTATCAAATTC 2580  
Db  
2521 GGCCAAATCTGATCCACTGAGATGCAATCTAGTAGAATCTCTCGCTATCAAATTC 2580  
Qy  
2581 ACTTCCACCTTCCACTCAGCGTTGTCATTCAGGCTGAACCTGCTTCTCTGTGAC 2640  
Db  
2581 ACTTCCACCTTCCACTCAGCGTTGTCATTCAGGCTGAACCTGCTTCTCTGTGAC 2640  
Qy  
2641 ATGACACACATCATCTCAATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGCCCA 2700  
Db  
2641 ATGACACACATCATCTCAATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGCCCA 2700  
Qy  
2701 TAAACACCAATAGCTTAACTATCCCAATTTTCCCAATTTTCCCAATTTCTCTTAATTTCA 2760  
Db  
2701 TAAACACCAATAGCTTAACTATCCCAATTTTCCCAATTTTCCCAATTTCTCTTAATTTCA 2760  
Qy  
2761 TGAACAACTCTTCAATCTTCTCTAGTCAATTTATTTGGTCCATTCACATTTCTCAT 2820  
Db  
2761 TGAACAACTCTTCAATCTTCTCTAGTCAATTTATTTGGTCCATTCACATTTCTCAT 2820  
Qy  
2821 CCCCTTTCCAGATTTAGATTTGCTTTCTTAATAAGATATTTGGAGAGCACCGTTC 2880  
Db  
2821 CCCCTTTCCAGATTTAGATTTGCTTTCTTAATAAGATATTTGGAGAGCACCGTTC 2880  
Qy  
2881 TTATTTCAGTATTAATAACTCGTCTTCTTAAGCATCCCTTCAATCTTCTTAATAACAATTA 2940  
Db  
2881 TTATTTCAGTATTAATAACTCGTCTTCTTAAGCATCCCTTCAATCTTCTTAATAACAATTA 2940  
Qy  
2941 TAGCATCTAATCTTCAACAACTGGCCGCTTGTGTGAACCTCTCTTTAATAAATAATTT 3000  
Db  
2941 TAGCATCTAATCTTCAACAACTGGCCGCTTGTGTGAACCTCTCTTTAATAAATAATTT 3000  
Qy  
3001 TTCCTTTCCCAATTTCCATTTAGATTTGCTTTCTTAATAAGATATTTGGAGAGCACCGTTC 3060  
Db  
3001 TTCCTTTCCCAATTTCCATTTAGATTTGCTTTCTTAATAAGATATTTGGAGAGCACCGTTC 3060  
Qy  
3061 TCACTGTATGAATCAAAATCGCTTCTTCTGTGTATCAAGGTTTAAATTTTATGTAAT 3120  
Db  
3061 TCACTGTATGAATCAAAATCGCTTCTTCTGTGTATCAAGGTTTAAATTTTATGTAAT 3120

Qy 3121 TCTTTTAAACAAACACCATAGAGATTAAACCTTTTACGGTGTAAACCTTCTCCCAATCA 3180  
Db 3121 TCTTTTAAACAAACACCATAGAGATTAAACCTTTTACGGTGTAAACCTTCTCCCAATCA 3180  
Qy 3181 GACAAACGTTTCAAATCTTTTCTTCATCATCGTGTATAAATCCGTATCCTTTACAGGA 3240  
Db 3181 GACAAACGTTTCAAATCTTTTCTTCATCATCGTGTATAAATCCGTATCCTTTACAGGA 3240  
Qy 3241 TATTTGAGTTCGTTCAATTCGTCGCGATTTAGTATATCCGATTTATATTTTTCGGTATT 3300  
Db 3241 TATTTGAGTTCGTTCAATTCGTCGCGATTTAGTATATCCGATTTATATTTTTCGGTATT 3300  
Qy 3301 TTTTATTTAAACGCTCTCAAAATCGTTTCTGGACGTTTCTAGCGTTTATTTTCGGTTAGTTA 3360  
Db 3301 TTTTATTTAAACGCTCTCAAAATCGTTTCTGGACGTTTCTAGCGTTTATTTTCGGTTAGTTA 3360  
Qy 3361 TCGGCATTAATCGTTTAAACAGCGTTATCGTAGCGTAAAGCCCTTGAGCGTAGCGTCT 3420  
Db 3361 TCGGCATTAATCGTTTAAACAGCGTTATCGTAGCGTAAAGCCCTTGAGCGTAGCGTCT 3420  
Qy 3421 TTGCAACGAGATGTTGTCGTGTAGATTATGAAGCCGATGACTGAATGAATAATAAAGC 3480  
Db 3421 TTGCAACGAGATGTTGTCGTGTAGATTATGAAGCCGATGACTGAATGAATAATAAAGC 3480  
Qy 3481 GCAGCGTCTCTTCTATTTTCGGTTTGGAGAGGCTCAAGGAGGTTTGAAGGAATGAATTC 3540  
Db 3481 GCAGCGTCTCTTCTATTTTCGGTTTGGAGAGGCTCAAGGAGGTTTGAAGGAATGAATTC 3540  
Qy 3541 TCATCGGTTTGAATTTTAAATTCCTTCAATTTTCCGAGCGGTAGCGCTGGAATTT 3600  
Db 3541 TCATCGGTTTGAATTTTAAATTTGCTTCAATTTTCCGAGCGGTAGCGCTGGAATTT 3600  
Qy 3601 TTGAAACAAATTTGGAATTTTGGAAAAAATGGGGGAAAGGAAGCGAATTTTTCCTCGT 3660  
Db 3601 TTGAAACAAATTTGGAATTTTGGAAAAAATGGGGGAAAGGAAGCGAATTTTTCCTCGT 3660  
Qy 3661 ACTACACCCCATTAAGTCCGAGTCCGAATTTTGTGCCAAAAAAGCTCTATCCCA 3720  
Db 3661 ACTACACCCCATTAAGTCCGAGTCCGAATTTTGTGCCAAAAAAGCTCTATCCCA 3720  
Qy 3721 CTGGCTCAAGGTTTGAAGGTTTTCATCCCAACGAATCCCAACGTTTTCGCCAAC 3780  
Db 3721 CTGGCTCAAGGTTTGAAGGTTTTCATCCCAACGAATCCCAACGTTTTCGCCAAC 3780  
Qy 3781 GTTTTTTATAAATCTATATTTAAGTAGCTTTATTTGTTTATGATTACAAAGTGATA 3840  
Db 3781 GTTTTTTATAAATCTATATTTAAGTAGCTTTATTTGTTTATGATTACAAAGTGATA 3840  
Qy 3841 CACTAATTTTATAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 3900  
Db 3841 CACTAATTTTATAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 3900  
Qy 3901 AAAGAGTTATGATTTCTCTGACAAAGAGCAAGATAAATAAATTAACAGATATGGCAAC 3960  
Db 3901 AAAGAGTTATGATTTCTCTGACAAAGAGCAAGATAAATAAATTAACAGATATGGCAAC 3960  
Qy 3961 AAAAGGTTTTCACAAATCTGCGTTGCGGTTAGCTATAGAGAAATTAAGCAAGAGG 4020  
Db 3961 AAAAGGTTTTCACAAATCTGCGTTGCGGTTAGCTATAGAGAAATTAAGCAAGAGG 4020  
Qy 4021 AATCAGAACAAAAAATAAGCGAAGCTCGGTTTGAAGGATACGAGTTTTCGCTA 4080  
Db 4021 AATCAGAACAAAAAATAAGCGAAGCTCGGTTTGAAGGATACGAGTTTTCGCTA 4080  
Qy 4081 CTTGTTTTTGAAGGTAATATATCATGCTATTAATAATACTAAGCTAGAAATTTTGG 4140  
Db 4081 CTTGTTTTTGAAGGTAATATATCATGCTATTAATAATACTAAGCTAGAAATTTTGG 4140  
Qy 4141 ATTTTTTATATATCTCTGACTCAATTCCTAATCGTGAAGAAATTAAGAGTTTGG 4200  
Db 4141 ATTTTTTATATATCTCTGACTCAATTCCTAATCGTGAAGAAATTAAGAGTTTGG 4200

4201 CGTATCTATGGCTGTGCTGCTTACACGATATGGACGAAAGAAAGATAAGATACATG 4260  
 4201 CGTATCTATGGCTGTGCTGCTTACACGATATGGACGAAAGAAAGATAAGATACATG 4260  
 4261 GAATAGTAGTGATGTTATACGAAATGAAAGCACTATAAAAAACCACTATCACGCTTAT 4320  
 4261 GAATAGTAGTGATGTTATACGAAATGAAAGCACTATAAAAAACCACTATCACGCTTAT 4320  
 4321 ATATATTGCGCGAAATCTCTAACAATAGAAAGCGTTAGGAAACAAGTTAAGCGAAATTT 4380  
 4321 ATATATTGCGCGAAATCTCTAACAATAGAAAGCGTTAGGAAACAAGTTAAGCGAAATTT 4380  
 4381 GGGGAATAGTTCAGTTCCTCATCTGATGATCTTGTATATATCAAAAGGTTTCATATCAATA 4440  
 4381 GGGGAATAGTTCAGTTCCTCATCTGATGATCTTGTATATATCAAAAGGTTTCATATCAATA 4440  
 4441 TTTGACTCATGAATCAAAAGGACGCTATTGCTAAGAAATAACATATATACGACAAAGAAAGA 4500  
 4441 TTTGACTCATGAATCAAAAGGACGCTATTGCTAAGAAATAACATATATACGACAAAGAAAGA 4500  
 4501 TATTTTGAACATTAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 4560  
 4501 TATTTTGAACATTAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 4560  
 4561 AAGAGAATTGGAAGAAATTTACTTTTAGATATAGTGGATGACTATAATTTGGTAAATACAAA 4620  
 4561 AAGAGAATTGGAAGAAATTTACTTTTAGATATAGTGGATGACTATAATTTGGTAAATACAAA 4620  
 4621 AGATTTAATGGCTTTTATTCGCTTAGGGAGCGGAGTTTGGAAATTTTAAATACGAATGA 4680  
 4621 AGATTTAATGGCTTTTATTCGCTTAGGGAGCGGAGTTTGGAAATTTTAAATACGAATGA 4680  
 4681 TGTAAAAGATATTGTTTCAACAACTCTAGCGCTTTAGATTTAGTTTGGGCGCAATTA 4740  
 4681 TGTAAAAGATATTGTTTCAACAACTCTAGCGCTTTAGATTTAGTTTGGGCGCAATTA 4740  
 4741 TCAGTGTGGATATAGCAAGTATATGCAAAAGTTCCTGATGCTGAAACGGGGGAAATAAA 4800  
 4741 TCAGTGTGGATATAGCAAGTATATGCAAAAGTTCCTGATGCTGAAACGGGGGAAATAAA 4800  
 4801 ATGCAACAAAGAAAGAGTATTGCTGAAATGAGGAAATAAAAGAAAGATTAAG 4860  
 4801 ATGCAACAAAGAAAGAGTATTGCTGAAATGAGGAAATAAAAGAAAGATTAAG 4860  
 4861 GACTTAAAGAGCGTATTGAAAGATACAGAGAAATGGAAGTTGAATTAAGTACAAATA 4920  
 4861 GACTTAAAGAGCGTATTGAAAGATACAGAGAAATGGAAGTTGAATTAAGTACAAATA 4920  
 4921 GATTTATGAGAGGAGGATATGATTAATAAAGCCCTGACGAAAGTCGGACT 4980  
 4921 GATTTATGAGAGGAGGATATGATTAATAAAGCCCTGACGAAAGTCGGACT 4980  
 4981 TCGTTCCTTTTACCTCTCGGTTATGAGTATGATTCAAATTCGTTCTTTTAGGTTCTAA 5040  
 4981 TCGTTCCTTTTACCTCTCGGTTATGAGTATGATTCAAATTCGTTCTTTTAGGTTCTAA 5040  
 5041 ATCGTGTTTTCTGGAAATGCTGTTTTATCGTTTACCTGCTACAAACCCCTTAA 5100  
 5041 ATCGTGTTTTCTGGAAATGCTGTTTTATCGTTTACCTGCTACAAACCCCTTAA 5100  
 5101 AACGTTTTTAAAGGCTTTTAAAGCGCTGCTGACGTTCTTAAAG 5142  
 5101 AACGTTTTTAAAGGCTTTTAAAGCGCTGCTGACGTTCTTAAAG 5142

RESULT 2  
 AR002333/c  
 LOCUS AR002333 8119 bp DNA linear PAT 04-DEC-1998  
 DEFINITION Sequence 1 from patent US 5741664.  
 ACCESSION AR002333  
 VERSION AR002333.1 GI:3963887  
 KEYWORDS Unknown.

ORGANISM Unknown.  
 Unclassified.  
 REFERENCE 1 (bases 1 to 8119)  
 AUTHORS Ballinger,M.D. and Wells,J.A.  
 TITLE Subtilisin variants capable of cleaving substrates containing  
 dibasic residues  
 JOURNAL Patent: US 5741664-A 1 21-APR-1998;  
 FEATURES  
 source Location/Qualifiers  
 1..8119  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 48.0%; Score 2467.2; DB 6; Length 8119;  
 Best local Similarity 99.7%; Pred. No. 0;  
 Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 831 TTGAAAAACCGACTGTGTAAGAGTACAGTCGGCATTTCTCATATATAAAGCCAGTCAT 890  
 DB 7847 TTGAAAAACCGACTGTGTAAGAGTACAGTCGGCATTTCTCATATATAAAGCCAGTCAT 7788  
 QY 891 TAGCCCTATCTGACAAATTCCTGATAGAGTTTCAAAAACATCTCGCATGATAACCATCAC 950  
 DB 7787 TAGCCCTATCTGACAAATTCCTGATAGAGTTTCAAAAACATCTCGCATGATAACCATCAC 7728  
 QY 951 AAACAGAAATGATGTACCTGTAAAGATAGCGGTAAATATATTTGAATTAACCTTTTAAATGA 1010  
 DB 7727 AAACAGAAATGATGTACCTGTAAAGATAGCGGTAAATATATTTGAATTAACCTTTTAAATGA 7668  
 QY 1011 ATTTTCTGCTGTAAATATGCGGTAGAGGTAATTTACTATTTATTATTTAGATTTTAAAGTTAA 1070  
 DB 7667 ATTTTCTGCTGTAAATATGCGGTAGAGGTAATTTACTATTTATTATTTAGATTTTAAAGTTAA 7608  
 QY 1071 ACCCAGTAAATGAACTCCATGGAATAATAGAAAGAGAAAAAGCATTTTCAGGTATAGGTG 1130  
 DB 7607 ACCCAGTAAATGAACTCCATGGAATAATAGAAAGAGAAAAAGCATTTTCAGGTATAGGTG 7548  
 QY 1131 TTTTGGGAAACAATTTCCCGAACCATTTATTTCTCTACATCAGAAAGGTATAAATCAT 1190  
 DB 7547 TTTTGGGAAACAATTTCCCGAACCATTTATTTCTCTACATCAGAAAGGTATAAATCAT 7488  
 QY 1191 AAAAATCTTTTCAAGTCACTTTTACAGAGTCCAAATACCAAGAGAAATGTTTTAGATACAC 1250  
 DB 7487 AAAAATCTTTTCAAGTCACTTTTACAGAGTCCAAATACCAAGAGAAATGTTTTAGATACAC 7428  
 QY 1251 CATCAAAAATTTGTAATAAGTGGCTCTAACTTATCCCAATAACCTAACTCTCCGTCGGTAT 1310  
 DB 7427 CATCAAAAATTTGTAATAAGTGGCTCTAACTTATCCCAATAACCTAACTCTCCGTCGGTAT 7368  
 QY 1311 TGTAAACAGTTCTAAAAGCTGTATTTGAGTTTATCACCTTGTACCTTAAGAAATAAATG 1370  
 DB 7367 TGTAAACAGTTCTAAAAGCTGTATTTGAGTTTATCACCTTGTACCTTAAGAAATAAATG 7308  
 QY 1371 CAGGTTAAAAATTTATATCTTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTT 1430  
 DB 7307 CAGGTTAAAAATTTATATCTTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTT 7248  
 QY 1431 CTGTGTTTATCTAAAAGTGGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1490  
 DB 7247 CTGTGTTTATCTAAAAGTGGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 7188  
 QY 1491 TCCAAATTTCTAAAATCAATTTTATTAAGTTCAATTTGATATGCTTCTCTCTCTCTCTCTCTCT 1550  
 DB 7187 TCCAAATTTCTAAAATCAATTTTATTAAGTTCAATTTGATATGCTTCTCTCTCTCTCTCTCTCT 7128  
 QY 1551 AAAGTGAATTTAGAGGCTTACTTGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1610  
 DB 7127 AAAGTGAATTTAGAGGCTTACTTGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 7068  
 QY 1611 AGTCAATATTTACTGTAAACATAAATATATATTTTAAAAAATATCCACATTTTATCCAAATTTTC 1670  
 DB 7067 AGTCAATATTTACTGTAAACATAAATATATATTTTAAAAAATATCCACATTTTATCCAAATTTTC 7008



QY 1671 GTTTGTTGAACTAATGGGTGCTTTAGTTGAAGATAAAGACCACATTAATAAAATGGTTC 1730  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5868  
QY 7007 GTTTGTTGAACTAATGGGTGCTTTAGTTGAAGATAAAGACCACATTAATAAAATGGTTC 6948  
QY 1731 TTTTGTGTTTTTTTAAAGATTTGAGCGTAGCGAATAATCCTTTCTTCTTCTTCTTCTGAT 1790  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5808  
QY 6947 TTTTGTGTTTTTTTAAAGATTTGAGCGTAGCGAATAATCCTTTCTTCTTCTTCTTCTGAT 6888  
QY 1791 AATAAGGGTAACATTTCCGGGATAGACTGTAACTCTCAGCATATAAATCCCTTTTCA 1850  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5748  
QY 6887 AATAAGGGTAACATTTCCGGGATAGACTGTAACTCTCAGCATATAAATCCCTTTTCA 6828  
QY 1851 TTTTCTAATGTAATCTATTACCTTATTAATTAATCAATTCGTCATTAATTAATTCCTTTT 1910  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5688  
QY 6827 TTTTCTAATGTAATCTATTACCTTATTAATTAATCAATTCGTCATTAATTAATTCCTTTT 6768  
QY 1911 TCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTATTCGTTAATTCGCCCATGAC 1970  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 3050  
QY 6767 TCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTATTCGTTAATTCGCCCATGAC 6708  
QY 1971 AGCCATGATAATTAATACTAGGAGAGTTAATAAATACGTAACCAACATGATTAACA 2030  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5628  
QY 6707 AGCCATGATAATTAATACTAGGAGAGTTAATAAATACGTAACCAACATGATTAACA 6648  
QY 2031 ATTAATTAGAGTCACTGTTCAAAATGGTATGCGTTTTGACACATCCACTATATATCGTG 2090  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 3110  
QY 6647 ATTAATTAGAGTCACTGTTCAAAATGGTATGCGTTTTGACACATCCACTATATATCGTG 6588  
QY 2091 TCGTTCTGTCACCTCCCTGATCCCAATTCAGAAATCTCTAGGAGATTCAGAAAGTTTCTC 2150  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5568  
QY 6587 TCGTTCTGTCACCTCCCTGATCCCAATTCAGAAATCTCTAGGAGATTCAGAAAGTTTCTC 6528  
QY 2151 AGAGTCGGAAGTTGACAGACATTAAGAACTGGCAGAGATGTCATTAACCTGAAGGAAG 2210  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 3230  
QY 6527 AGAGTCGGAAGTTGACAGACATTAAGAACTGGCAGAGATGTCATTAACCTGAAGGAAG 6468  
QY 2211 ATCTGATTGCTTAATCTGTTTCAAGTAAGACCGAGCGCTGTCGTATAACAGATGCGATG 2270  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5448  
QY 6467 ATCTGATTGCTTAATCTGTTTCAAGTAAGACCGAGCGCTGTCGTATAACAGATGCGATG 6408  
QY 2271 ATGAGACCAATCAACATGCGACCTGCCATTTGCTTACCTGTACAGTCAAGATGGTAGAAA 2330  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 3290  
QY 6407 ATGAGACCAATCAACATGCGACCTGCCATTTGCTTACCTGTACAGTCAAGATGGTAGAAA 6348  
QY 2331 TGTGTCGTCCTTGCACAGAAATTAAGCCATTTACGCCATTTCCCTGCAATTAACAAGCTCTTC 2390  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5448  
QY 6347 TGTGTCGTCCTTGCACAGAAATTAAGCCATTTACGCCATTTCCCTGCAATTAACAAGCTCTTC 6288  
QY 2391 TACGATAAGGCGACAAATCGCATCGTGAAGCTTTGGGCTTTCTACCGATTTAGCAGTTG 2450  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 3290  
QY 6287 TACGATAAGGCGACAAATCGCATCGTGAAGCTTTGGGCTTTCTACCGATTTAGCAGTTG 6228  
QY 2451 ATACACTTCTTAAGTATCCACTGAATCATTAATCGGCAATCGGCAATAGAGAAATTCACC 2510  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5448  
QY 6227 ATACACTTCTTAAGTATCCACTGAATCATTAATCGGCAATCGGCAATAGAGAAATTCACC 6168  
QY 2511 ATGTGTAAGCGGCAATCTGATTCACCTGAGATGCATTAATCTAGTAGAATCTCTCGCT 2570  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5448  
QY 6167 ATGTGTAAGCGGCAATCTGATTCACCTGAGATGCATTAATCTAGTAGAATCTCTCGCT 6108  
QY 2571 ATCAAAATTCATCTCCACTTCCACTCACCGGTTGTCATTCATCGGTGAACTCTGCTTC 2630  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5448  
QY 6107 ATCAAAATTCATCTCCACTTCCACTCACCGGTTGTCATTCATCGGTGAACTCTGCTTC 6048  
QY 2631 CTCTGTTGACATGACACATCATCTCAATTCGMAATAGGCGCCATCAGTCTCAGCACC 2690  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5448  
QY 6047 CTCTGTTGACATGACACATCATCTCAATTCGMAATAGGCGCCATCAGTCTCAGCACC 5988  
QY 2691 AAGAGAGCCATAAACCACCAATAGGCTTAACTATCCCAATTTATCCCAATATTCGTTTC 2750  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5448  
QY 5987 AAGAGAGCCATAAACCACCAATAGGCTTAACTATCCCAATTTATCCCAATATTCGTTTC 5928  
QY 2751 CTTAAATTCATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 2810

Db 5927 CTTAAATTCATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5868  
QY 2811 TATTCCTCATCTCCTTTTCAGATAATTTTAGAATTTCTTTCTTAAATAAGAAATATTGGAG 2870  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5808  
QY 5867 TATTCCTCATCTCCTTTTCAGATAATTTTAGAATTTCTTTCTTAAATAAGAAATATTGGAG 5808  
QY 2871 AGCACCGTTCTTATTTCAGCTATTAATAACTCGTCTTCTTAAGCATCTTCAATCCCTTTTA 2930  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5748  
QY 5807 AGCACCGTTCTTATTTCAGCTATTAATAACTCGTCTTCTTAAGCATCTTCAATCCCTTTTA 5748  
QY 2931 ATAAACAATATAGCATCTAAATCTTCAACAACATGCGCGTTTGTGTAACCTACTCTTTAAT 2990  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5688  
QY 5747 ATAAACAATATAGCATCTAAATCTTCAACAACATGCGCGTTTGTGTAACCTACTCTTTAAT 5688  
QY 2991 AAAATAAATTTTTCGTTTCCCAATTCACATTCGCAATATAGAAAAATCATCTTCAATCGC 3050  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5628  
QY 5687 AAAATAAATTTTTCGTTTCCCAATTCGCAATTCGCAATATAGAAAAATCATCTTCAATCGC 5628  
QY 3051 TTTTTCGTCATCATCTGTATGATCAAAATCGCTTCTTCTGTGTCATCAAGGTTTAAATTT 3110  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5568  
QY 5627 TTTTTCGTCATCATCTGTATGATCAAAATCGCTTCTTCTGTGTCATCAAGGTTTAAATTT 5568  
QY 3111 TTTATGATTTCTTTTAAACAAACCATAGAGATTAACCTTTTACGGTGTAAACCTTC 3170  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5508  
QY 5567 TTTATGATTTCTTTTAAACAAACCATAGAGATTAACCTTTTACGGTGTAAACCTTC 5508  
QY 3171 CTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCTCATCATCGGTCAATAAATCCGTATC 3230  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5448  
QY 5507 CTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCTCATCATCGGTCAATAAATCCGTATC 5448  
QY 3231 CTTTACAGATATTTTGCAGTTTCTCAATTCGCGATTTGATATCCGATTTTATTTATT 3290  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5388  
QY 5447 CTTTACAGATATTTTGCAGTTTCTCAATTCGCGATTTGATATCCGATTTTATTTATT 5388  
QY 3291 TTTTCGTTATTTTATTATAA 3310  
Db |||||||CATGAACAATCTTCTTCTCTAGTCATTAATTTGGTCCATTAC 5387

RESULT 3  
AR018032/C

LOCUS 8119 bp DNA linear PAT 05-DEC-1998  
DEFINITION Sequence 1 from patent US 5780285.  
ACCESSION AR018032  
VERSION AR018032.1 GI:3973635  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 8119)  
AUTHORS Ballinger, M.D. and Wells, J.A.  
TITLE Subtilisin variants capable of cleaving substrates containing  
dibasic residues  
JOURNAL Patent: US 5780285-A 1 14-JUL-1998;  
FEATURES Location/Qualifiers  
source 1. 8119  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 48.0%; Score 2467.2; DB 6; Length 8119;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 831 TTAGAAAACCGACTGTAAAGATACAGTCGGCATTTATCTCATATTATAAAGCCAGTCAT 890  
Db 7847 TTAGAAAACCGACTGTAAAGATACAGTCGGCATTTATCTCATATTATAAAGCCAGTCAT 7788  
QY 891 TAGGCGTATCTGAATTCCTGAAATAGAGTTTCATAAACAATCCTGTCATGATCAACATCAC 950  
Db 7787 TAGGCGTATCTGAATTCCTGAAATAGAGTTTCATAAACAATCCTGTCATGATCAACATCAC 7728

Qy	951	AAACAGAAATGATGTA	CTGCTAAAGATAGCGGTAAATATATTTGAATTAACCTTTATTAATGA	1010
Db	7727	AAACAGAAATGATGTA	CTGCTAAAGATAGCGGTAAATATATTTGAATTAACCTTTATTAATGA	7668
Qy	1011	ATTTTCCTGCTGTAAT	TAATCGGTAGAGGTAAATTA	1070
Db	7667	ATTTTCCTGCTGTAAT	TAATCGGTAGAGGTAAATTA	7608
Qy	1071	ACCAGTAAATGAAGT	CCATCGAATTAATAGAAAGAGAAAAGCATTTTCAGGTATAGGTG	1130
Db	7607	ACCAGTAAATGAAGT	CCATCGAATTAATAGAAAGAGAAAAGCATTTTCAGGTATAGGTG	7548
Qy	1131	TTTTGGGAAAACAAT	TTCCCGCAACCATTAATATTTCTACATCAGAAAGGTATAAATCAT	1190
Db	7547	TTTTGGGAAAACAAT	TTCCCGCAACCATTAATATTTCTACATCAGAAAGGTATAAATCAT	7488
Qy	1191	AAAACCTTTTGAAGT	CAATCTTTACAGAGTCCAAATACAGAGAAATGTTTTAGATACAC	1250
Db	7487	AAAACCTTTTGAAGT	CAATCTTTACAGAGTCCAAATACAGAGAAATGTTTTAGATACAC	7428
Qy	1251	CATCAAAAAATGTAT	ATAAGTGGCTTAACCTTATCCCAATACCTTA	1310
Db	7427	CATCAAAAAATGTAT	ATAAGTGGCTTAACCTTATCCCAATACCTTA	7368
Qy	1311	TGTAACCAAGTTC	TAAAGCTGTATTTGAGTTTATCACCCCTTGTCACCTAAGAAAATAAATG	1370
Db	7367	TGTAACCAAGTTC	TAAAGCTGTATTTGAGTTTATCACCCCTTGTCACCTAAGAAAATAAATG	7308
Qy	1371	CAGGGTAAAAATTTAT	ATPCTTCCTGTTTTATGTTTGGTATAAAAACATAAATCAATTT	1430
Db	7307	CAGGGTAAAAATTTAT	ATPCTTCCTGTTTTATGTTTGGTATAAAAACATAAATCAATTT	7248
Qy	1431	CTGTGGTTATAC	TAAAGAGTCGTTGTTGGTTTCAAAATAATGATTAATAATCTCTTTTCTCT	1490
Db	7247	CTGTGGTTATAC	TAAAGAGTCGTTGTTGGTTTCAAAATAATGATTAATAATCTCTTTTCTCT	7188
Qy	1491	TCCAATTTGCTCTAAAT	CAATTTTATTAAGTTTCATTTTGATATGCTCTAAATTTTATCT	1550
Db	7187	TCCAATTTGCTCTAAAT	CAATTTTATTAAGTTTCATTTTGATATGCTCTAAATTTTATCT	7128
Qy	1551	AAAGTGAATTTAGAGAGCTT	ACTTGCTGCTTTTCATTAAGATCAATCTCTTTTATAA	1610
Db	7127	AAAGTGAATTTAGAGAGCTT	ACTTGCTGCTTTTCATTAAGATCAATCTCTTTTATAA	7068
Qy	1611	AGTCAATATTACTGTAA	CAATAAATATATATTTTAAAAATATCCACATTTATCCAATTTTC	1670
Db	7067	AGTCAATATTACTGTAA	CAATAAATATATATTTTAAAAATATCCACATTTATCCAATTTTC	7008
Qy	1671	GTTTGTGTTGAATTAAT	GGGTGCTTTAGTTGAAGATAAAGACCAATTA	1730
Db	7007	GTTTGTGTTGAATTAAT	GGGTGCTTTAGTTGAAGATAAAGACCAATTA	6948
Qy	1731	TTTTGTGTTTTTTTTTAA	AGGAATTCAGCGGTAGCGAAAAATCCTTTCTTTCTTATCTTGAT	1790
Db	6947	TTTTGTGTTTTTTTTTAA	AGGAATTCAGCGGTAGCGAAAAATCCTTTCTTTCTTATCTTGAT	6888
Qy	1791	AATAGGGTAACTAT	ATGCGGGATAGCTGTAACTTCTCAGCATAAAAATCCCTTTTCA	1850
Db	6887	AATAGGGTAACTAT	ATGCGGGATAGCTGTAACTTCTCAGCATAAAAATCCCTTTTCA	6828
Qy	1851	TTTTCTAATGTAAT	CTATTATACCTTATTTAAATCAATTCGCTCAATAATAATCCCTTTT	1910
Db	6827	TTTTCTAATGTAAT	CTATTATACCTTATTTAAATCAATTCGCTCAATAATAATCCCTTTT	6768
Qy	1911	TCTTATTACGAAAAAT	GCGCATTTTAAGCACACCCCTTTATTCGGTTAATTCGGCATGAC	1970
Db	6767	TCTTATTACGAAAAAT	GCGCATTTTAAGCACACCCCTTTATTCGGTTAATTCGGCATGAC	6708
Qy	1971	AGCCATGATTAAT	TACTAATCTAGAGAGTTAATAATACGTAAACCAACATGATTAACA	2030
Db	6707	AGCCATGATTAAT	TACTAATCTAGAGAGTTAATAATACGTAAACCAACATGATTAACA	6648
Qy	2031	ATTATTAGAGTCACT	CGTTCAAAATGTTATGCGTTTTTGACACATCCACTATATATCCGTG	2090

Db	6647	ATTATTAGAGTCATCGTTCAAAATGGTATGGTTTTGACACATCCACTATATATCCGTG	6598
Qy	2091	TCGTTCTGTCCACTCCTGGAATCCCAATTCAGAAAATTCCTTAGCGAATCCAGAACTTCTC	2150
Db	6587	TCGTTCTGTCCACTCCTGGAATCCCAATTCAGAAAATTCCTTAGCGAATCCAGAACTTCTC	6528
Qy	2151	AGATCGGAAGTTGACCGACACATTACGAACTGGCACAGATGCTCATAACTGAAAGGAAG	2210
Db	6527	AGATCGGAAGTTGACCGACACATTACGAACTGGCACAGATGCTCATAACTGAAAGGAAG	6468
Qy	2211	ATCTGATTGCTTAACTGCTTCAGTTAAGACCGAAGCGCTCGTCGPTATTAACAGATCGCATG	2270
Db	6467	ATCTGATTGCTTAACTGCTTCAGTTAAGACCGAAGCGCTCGTCGPTATTAACAGATCGCATG	6408
Qy	2271	ATGCAGACCAATCAACATGCGCACCTGCCATTGCTACTGTTACAGTCAAGGATGTTAGAAA	2330
Db	6407	ATGCAGACCAATCAACATGCGCACCTGCCATTGCTACTGTTACAGTCAAGGATGTTAGAAA	6348
Qy	2331	TGTTGTGCGTCTTTGCACACGAATAATTAGCCATTGTCCTGCTATATTCAAAACAGCTCTTC	2390
Db	6347	TGTTGTGCGTCTTTGCACACGAATAATTAGCCATTGTCCTGCTATATTCAAAACAGCTCTTC	6288
Qy	2391	TACGATAAGGCGCAAAATCGATCGTGGAAAGCTTTGGGCTTCTACCGATTTAGCAGTTTG	2450
Db	6287	TACGATAAGGCGCAAAATCGATCGTGGAAAGCTTTGGGCTTCTACCGATTTAGCAGTTTG	6228
Qy	2451	ATACACTTTCTTAAGTATCCACTCAATATATAAATCGGCAAAATAGAGAAAAATTGACC	2510
Db	6227	ATACACTTTCTTAAGTATCCACTCAATATATAAATCGGCAAAATAGAGAAAAATTGACC	6168
Qy	2511	ATGTGTAGGGCGGCGAATCTGATTTCGACCTGAGATGCAATCTAGTAGAATCTCTTCGCT	2570
Db	6167	ATGTGTAGGGCGGCGAATCTGATTTCGACCTGAGATGCAATCTAGTAGAATCTCTTCGCT	6108
Qy	2571	ATCAAAATTCACATCCACCTTCCACTCCCGGTTGTCATTCATGGCTGAACTCTGCTTC	2630
Db	6107	ATCAAAATTCACATCCACCTTCCACTCCCGGTTGTCATTCATGGCTGAACTCTGCTTC	6048
Qy	2631	CTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGCCCATCAGTCTGACGACC	2690
Db	6047	CTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGCCCATCAGTCTGACGACC	5988
Qy	2691	AAGAGAGCCATAACACCAATAGCCTTACATCATCCCATTTTATCCAAATATTCGTTTC	2750
Db	5987	AAGAGAGCCATAACACCAATAGCCTTACATCATCCCATTTTATCCAAATATTCGTTTC	5928
Qy	2751	CTTAATTTTCATGAACAAATCTTCATTCTTCTCTAGTCATTATTAATGGTCCATTCCAC	2810
Db	5927	CTTAATTTTCATGAACAAATCTTCATTCTTCTCTAGTCATTATTAATGGTCCATTCCAC	5868
Qy	2811	TATTCCTGATCCCTTTTCAGATAATTTAGATTTGCTTTTCTAAATAGAAATATTTGGAG	2870
Db	5867	TATTCCTGATCCCTTTTCAGATAATTTAGATTTGCTTTTCTAAATAGAAATATTTGGAG	5808
Qy	2871	AGCACCGTCTTATTCAGCTATTAAATACTCGTCTTCCTAAGCATCCCTTCAATCCTTTTA	2930
Db	5807	AGCACCGTCTTATTCAGCTATTAAATACTCGTCTTCCTAAGCATCCCTTCAATCCTTTTA	5748
Qy	2931	ATAACAAATATAGCAATCTAATCTTCAACAAACTGGCCCGTTTGTGTGAACACTCTTTAAT	2990
Db	5747	ATAACAAATATAGCAATCTAATCTTCAACAAACTGGCCCGTTTGTGTGAACACTCTTTAAT	5688
Qy	2991	AAAAATAATTTTTCGGTTCCTCAATTCACATTTGCAATATAGAAAAATCCATCTTCATCGGC	3050
Db	5687	AAAAATAATTTTTCGGTTCCTCAATTCACATTTGCAATATAGAAAAATCCATCTTCATCGGC	5628
Qy	3051	TTTTTCGTGCATCATCTGTATGAATCAAAATCGCCTCTTCTGTGTGTCATCAGGTTTAATTT	3110
Db	5627	TTTTTCGTGCATCATCTGTATGAATCAAAATCGCCTCTTCTGTGTGTCATCAGGTTTAATTT	5568
Qy	3111	TTTTATGTATTTCTTTTAAACAAACCCACATAGGAGATTAACTTTTACGGTGTAAACCTTC	3170

Db 5567 TTTATGTAATCTCTTTTAAACAAACACATAGGAGATTAAACCTTTTACGGTGTAAACCTTC 5508  
Qy 3171 CTCCAATCAGACAAAGCTTTCAAAATCTTTCTTCATCATCGGTGATATAAAATCCGTATC 3230  
Db 5507 CTCCAATCAGACAAAGCTTTCAAAATCTTTCTTCATCATCGGTGATATAAAATCCGTATC 5448  
Qy 3231 CTTTACAGGATATTTTCAGATTTCGTCATTCGCAATTCGGGATTTGATATCCGATTTATTTAT 3290  
Db 5447 CTTTACAGGATATTTTCAGATTTCGTCATTCGCAATTCGGGATTTGATATCCGATTTATTTAT 5388  
Qy 3291 TTTTCGGTATTTTATTTAAA 3310  
Db 5387 TTTTCGGTCAATCATTTGAA 5368

RESULT 4  
AR055698/c  
LOCUS AR055698 8119 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5837516.  
ACCESSION AR055698  
VERSION AR055698.1 GI:5981275  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 8119)  
AUTHORS Ballinger, M.D. and Wells, J.A.  
TITLE Subtilisin variants capable of cleaving substrates containing basic residues  
JOURNAL Patent: US 5837516-A 1 17-NOV-1998;  
FEATURES Location/Qualifiers  
source 1..8119  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 48.0%; Score 2467.2; DB 6; Length 8119;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 831 TTGAAAACCCGACGTAAAGATGACGTCCGCTATCTCATATATAAAGCCAGTCAT 890  
Db 7847 TTGAAAACCCGACGTAAAGATGACGTCCGCTATCTCATATATAAAGCCAGTCAT 7788  
Qy 891 TAGCCCTATCTGACAAATTCCTGAATAGAGTTCAATAAATCCCTGCGATGATAACCATCAC 950  
Db 7787 TAGCCCTATCTGACAAATTCCTGAATAGAGTTCAATAAATCCCTGCGATGATAACCATCAC 7728  
Qy 951 AAACAGATGATGATACCTGTTAAAGATAGCGGTAAATATTTGAATTTACCTTTTATATGA 1010  
Db 7727 AAACAGATGATGATACCTGTTAAAGATAGCGGTAAATATTTGAATTTACCTTTTATATGA 7668  
Qy 1011 ATTTTCCTGCTGTAATATGGTGAAGGTAATTTACTATTATTATTGATATTTAAAGTTAA 1070  
Db 7667 ATTTTCCTGCTGTAATATGGTGAAGGTAATTTACTATTATTATTGATATTTAAAGTTAA 7608  
Qy 1071 ACCAGTAAATGAAGTCATCGAATAATAGAAAGAGAAAAGCAATTTTCAGGTATAGGTG 1130  
Db 7607 ACCAGTAAATGAAGTCATCGAATAATAGAAAGAGAAAAGCAATTTTCAGGTATAGGTG 7548  
Qy 1131 TTTTGGGAAACAAATTTCCCGAACCATATATTTCTCTACATCAGAAAGGTATAAATCAT 1190  
Db 7547 TTTTGGGAAACAAATTTCCCGAACCATATATTTCTCTACATCAGAAAGGTATAAATCAT 7488  
Qy 1191 AAAACTCTTTGAAGTCATTTTACAGGAGTCCAAATACAGAGAAATGTTTTAGATACAC 1250  
Db 7487 AAAACTCTTTGAAGTCATTTTACAGGAGTCCAAATACAGAGAAATGTTTTAGATACAC 7428  
Qy 1251 CATCAAAATTTGATAAAGTGGCTCTAATCTTATCCCAATACCTTAACTCCCGTCGTAT 1310  
Db 7427 CATCAAAATTTGATAAAGTGGCTCTAATCTTATCCCAATACCTTAACTCCCGTCGTAT 7368  
Qy 1311 TGTAAACCAGTTCTAAAGCTGTATTTGAGTTTATCACCCCTTGTCACTAAGAAAAATAATG 1370

Db 7367 TGTAAACCAGTTCTAAAGCTGTATTTGAGTTTATCACCCCTTGTCACTAAGAAATAATG 7308  
Qy 1371 CAGGGTAAAAATTTATATCTCTTCTGTTTATGTTTCGGTATATAAACACTAATATCAATTT 1430  
Db 7307 CAGGGTAAAAATTTATATCTCTTCTGTTTATGTTTCGGTATATAAACACTAATATCAATTT 7248  
Qy 1431 CTGTGGTTATATAAAGCTCGTTTGTGTTCAAAATATGATTAATAATCTCTTTTCTCT 1490  
Db 7247 CTGTGGTTATATAAAGCTCGTTTGTGTTCAAAATATGATTAATAATCTCTTTTCTCT 7188  
Qy 1491 TCCAAATGTCTAAATCAATTTTATTAAGTTTCAITTTGATATGCTCTCCPAATTTTATCT 1550  
Db 7187 TCCAAATGTCTAAATCAATTTTATTAAGTTTCAITTTGATATGCTCTCCPAATTTTATCT 7128  
Qy 1551 AAAGTGAATTTAGGAGGCTTACTGTCTGCTTTCTTCAATAGATCAATCTTTTAAA 1610  
Db 7127 AAAGTGAATTTAGGAGGCTTACTGTCTGCTTTCTTCAATAGATCAATCTTTTAAA 7068  
Qy 1611 AGTCAATATTTACTGTAAACATAAATAATATTTTAAAAATATCCCACTTTATCCAAATTTTC 1670  
Db 7067 AGTCAATATTTACTGTAAACATAAATAATATTTTAAAAATATCCCACTTTATCCAAATTTTC 7008  
Qy 1671 GTTGTGTAACATAATGGGTGCTTTAGTTGAGAATAAAGACACATTTAAAAATGCTGTC 1730  
Db 7007 GTTGTGTAACATAATGGGTGCTTTAGTTGAGAATAAAGACACATTTAAAAATGCTGTC 6948  
Qy 1731 TTTTGTGTTTTTTTAAAGGATTTGAGCTAGCGAAAAATCCCTTTCTTTCTTATCTTGAT 1790  
Db 6947 TTTTGTGTTTTTTTAAAGGATTTGAGCTAGCGAAAAATCCCTTTCTTTCTTATCTTGAT 6888  
Qy 1791 AATAAGGTPAATATTGCGGGATAGACTGTAAACATTTCTACGGCATAAAAATCCCTTTCA 1850  
Db 6887 AATAAGGTPAATATTGCGGGATAGACTGTAAACATTTCTACGGCATAAAAATCCCTTTCA 6828  
Qy 1851 TTTTCTAATGTAATCTATTACCTTTATTTAATTTCAATTCGCTCATATTAATCTCTTT 1910  
Db 6827 TTTTCTAATGTAATCTATTACCTTTATTTAATTTCAATTCGCTCATATTAATCTCTTT 6768  
Qy 1911 TCTTATTACGCAAAATGCGCCGATTTAAGCACACACCTTTTATTCGGTTAATGCGCATGAC 1970  
Db 6767 TCTTATTACGCAAAATGCGCCGATTTAAGCACACACCTTTTATTCGGTTAATGCGCATGAC 6708  
Qy 1971 AGCCATGATAATCTAATATCTAGGAGAGTTAATAATAGTAACCAACATGATTAACA 2030  
Db 6707 AGCCATGATAATCTAATATCTAGGAGAGTTAATAATAGTAACCAACATGATTAACA 6648  
Qy 2031 ATTATTAGAGGTCATCGTTTCAAAATGGTATGGTTTTGACACATCCACTATATATCCGTG 2090  
Db 6647 ATTATTAGAGGTCATCGTTTCAAAATGGTATGGTTTTGACACATCCACTATATATCCGTG 6588  
Qy 2091 TCGTTCTGTCCTCTGAATCCCATCCAGAAATTTCTAGCGATTCAGAAAGTTTCTC 2150  
Db 6587 TCGTTCTGTCCTCTGAATCCCATCCAGAAATTTCTAGCGATTCAGAAAGTTTCTC 6528  
Qy 2151 AGAGTCGGAAGTTGACACAGACATTTACGACCTGGCACAGATGGTTCATACCTGAAGGAG 2210  
Db 6527 AGAGTCGGAAGTTGACACAGACATTTACGACCTGGCACAGATGGTTCATACCTGAAGGAG 6468  
Qy 2211 ATCTGATTCCTTAACTGCTTTCAGTTTAAAGCCGAAGCGCTCGTCGTATTAACAGATCGATG 2270  
Db 6467 ATCTGATTCCTTAACTGCTTTCAGTTTAAAGCCGAAGCGCTCGTCGTATTAACAGATCGATG 6408  
Qy 2271 ATGCAGACCAATCAACATGGACCTGCCATTCCTACCTGTACAGTCAGGATGGTAGAA 2330  
Db 6407 ATGCAGACCAATCAACATGGACCTGCCATTCCTACCTGTACAGTCAGGATGGTAGAA 6348  
Qy 2331 TGTGTGCGTCTTGACACAGCAATATTACGCCATTTGCTGTCATATTCAAACAGCTCTTC 2390  
Db 6347 TGTGTGCGTCTTGACACAGCAATATTACGCCATTTGCTGTCATATTCAAACAGCTCTTC 6288  
Qy 2391 TACGATPAGGGCACAAATCGCATCGTGGGAACTGTTGGGCTTCTACCGATTTAGCAGTTG 2450







CDS	complement (5365..6063) /gene="repA" /codon_start=1 /transl_table=11 /product="replication protein RepA" /protein_id="AAD46053.1" /db_xref="GI:5668892" /translation="MAIKNTKARNFGLYPSIPNDWKEKLESIGVMAVSLPHDMD EKDKDTWNSDVIRNGHYKPHYHVIYIARNPVITSVNRKIRKLGNSVAHVBI LDYIKGSYELTHESDAJAKNKHIDYKDIININDFDIDRYITLDESOKRELKULL DIVDDYNVNTKDLMAFIRLGAEFGLINTNDYKDIVSTVNSSAFRLWFEQNYCCGYRA SYAKVLDARTGEIK" 7156..7215 /codon_start=1 /transl_table=11 /product="emrC leader peptide" /protein_id="AAD46054.1" /db_xref="GI:5668893" /translation="MGIFSIFVISTVHYQPNKK" 7275..8009 /gene="emrC" 7275..8009 /gene="emrC" /codon_start=1 /transl_table=11 /product="adenine methylase" /protein_id="AAD46055.1" /db_xref="GI:5668894" /translation="MNEKNIKHSQNFITSKHNDIKIMWIRLNEHNDIPEISGKGFH TLELVKCNFVTAIEDHLKCTTENKLVDFQVINKDIDWQFPPKQNSYKIYGN EYHSTDIIRKIVPDSIANEYIYIVGYFAKRLNTRSLALLLMAEVDISLSMVP EYHPKPNVNSLIRLSRKSRISHKDKQKYNFVMKWNKEYKFIITKNQFNLSLKH AGIDLLNISFEQFLSLFNSYKLFNK"	
ORIGIN		
Query Match	31.9%; Score 1641.6; DB 12; Length 8082;	
Best Local Similarity	98.4%; Pred. No. 1e-263;	
Matches 1689; Conservative	0; Mismatches 24; Indels 3; Gaps 3;	
QY	3279	TTTATATTTATTTTCGGTATTTTATTAAGAGTCTCAAAATCGTTTCTGGGACGTTT 3338
DB	6892	TATATATTTATTTATCCGATTTTATTAAGAGTCTCAAAATCGTTTCTGGGACGTTT 6833
QY	3339	TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTAAACACGCGCTATCGTAGCGTAA 3398
DB	6832	TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTAAACACGCGCTATCGTAGCGTAA 6773
QY	3399	AAGCCCTTGACGCTAGCGT-GCTTTGACGGAAGATGTTGCTGTAGATTATGAAGCC 3457
DB	6772	AAGCCCTTGACGCTAGCGTGGCTTTGACGGAAGATGTTGCTGTAGATTATGAAGCC 6713
QY	3458	GATGACTGAATGAATAATGAAGCAGCGTCTTCTATTTTCGGTTGGAGGAGGCTCAAGG 3517
DB	6712	GATGACTGAATGAATAATGAAGCAGCGCTTCTATTTTCGGTTGGAGGAGGCTCAAGG 6653
QY	3518	CAGTTTGAGGAATGAATTCCTCATCGGGTTGATTTTAAATTCCTTGCATTTTGC 3577
DB	6652	CAGTTTGAGGAATGAATTCCTCATCGGGTTGATTTTAAATTCCTTGCATTTTGC 6593
QY	3578	CGAGCGGTAGCGCTGGAATTTTGGAAAAATTTTGGAAATTTTGGAAAAATTTGGGGGA 3637
DB	6592	CGAGCGGTAGCGCTGGAATTTTGGAAAAATTTTGGAAATTTTGGAAAAATTTGGGGGA 6533
QY	3638	AAGGAAGCGAATTTTGTCTTCGCTACTACGACCCCCCAATTAAGTCCCGAGTGCCAAATTTT 3697
DB	6532	AAGGAAGCGAATTTTGTCTTCGCTACTACGACCCCCCAATTAAGTCCCGAGTGCCAAATTTT 6473
QY	3698	GTGCCAAAAACGCTCTATCCCACTGCGCTCAAGGGTTGAGGGTTTTCATTCGCCAAC 3757
DB	6472	GTGCCAAAAACGCTCTATCCCACTGCGCTCAAGGGTTTTCATTCGCCAAC 6413
QY	3758	GAATCGCCAAAGCTTTTCGCCAACGTTTATTATAATCTATATTTAAGTAGCTTTATTGTT 3817

DB	6412	GAATCGCCCAACGTTTTCGCCCAACGTTTTTTATAAATCTATATTTAAGTAGCTTTATTGTT 6353
QY	3818	GTTTTTATGATTACAAAGTGATACACTAAATTTTATAAATTTATTTGATTGAGTTTATTA 3877
DB	6352	GTTTTTATGATTACAAAGTGATACACTAAATTTTATAAATTTATTTGATTGAGTTTATTA 6293
QY	3878	AATGTGATTTTCAGAAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 3937
DB	6292	AATGTGATTTTCAGAAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 6233
QY	3938	AAAAATACAGATATGCGGAAACAAAAGGTTTTTCAAAATCTCGGTTGCGGCTTAGC 3997
DB	6232	AAAAATACAGATATGCGGAAACAAAAGGTTTTTCAAAATCTCGGTTGCGGCTTAGC 6173
QY	3998	TATGAACAATATGCAAGAAAGGAATCAGAAACAAAATAAAGCGAAAGCTCGCGTTT 4057
DB	6172	TATGAACAATATGCAAGAAAGGAATCAGAAACAAAATAAAGCGAAAGCTCGCGTTT 6113
QY	4058	TAGAAGGATACGAGTTTTCGCTACTGTTTGTGTTTGAAGGTAA-TATATCATGGCTATTAA 4116
DB	6112	TAGAAGGATACGAGTTTTCGCTACTGTTTGTGTTTGAAGGTAA-TATATCATGGCTATTAA 6053
QY	4117	AAATACCTAAAGCTAGAAATTTTGGATTTTATTTATATCTCTGACCTCAATTCCTATGATTG 4176
DB	6052	AAATACCTAAAGCTAGAAATTTTGGATTTTATTTATATCTCTGACCTCAATTCCTATGATTG 5993
QY	4177	GAAGAAAAATTAGAGAGTTTGGGCGTATCTATGCTGTCTGCTCTTACACGATATGGA 4236
DB	5992	GAAGAAAAATTAGAGAGTTTGGGCGTATCTATGCTGTCTGCTCTTACACGATATGGA 5933
QY	4237	CGAAAAAAGATAAAGATACATGGAATAGTAGTGTATTTACGAAATGGAAGCACTA 4296
DB	5932	CGAAAAAAGATAAAGATACATGGAATAGTAGTGTATTTACGAAATGGAAGCACTA 5873
QY	4297	TAAAAAACCACTATACGTTTATATATTTGACGAAATCTGTAAACAATAGAAAGCGT 4356
DB	5872	TAAAAAACCACTATACGTTTATATATTTGACGAAATCTGTAAACAATAGAAAGCGT 5813
QY	4357	TAGGAACAAGATTAAAGCGAAAAATTTGGGAAATAGTTTCACTGTCTCATTTGAGATACCTGA 4416
DB	5812	TAGGAACAAGATTAAAGCGAAAAATTTGGGAAATAGTTTCACTGTCTCATTTGAGATACCTGA 5753
QY	4417	TTATATCAAGGTTTCATATGATATTTGACTCATGAAATCAAGAGGACGCTATTGCTAAGAA 4476
DB	5752	TTATATCAAGGTTTCATATGATATTTGACTCATGAAATCAAGAGGACGCTATTGCTAAGAA 5693
QY	4477	TAAACATATATACGACAAAAAGATATTTTGAACATTAATGATTTTGAATTTGATACCGCTA 4536
DB	5692	TAAACATATATACGACAAAAAGATATTTTGAACATTAATGATTTTGAATTTGATACCGCTA 5633
QY	4537	TATAACACTTGTAGAAAGCCAAAAAGAGATTTGAAGATTTTACTTTTAGATATAGTGA 4596
DB	5632	TATAACACTTGTAGAAAGCCAAAAAGAGATTTGAAGATTTTACTTTTAGATATAGTGA 5573
QY	4597	TGACTATATTTTGGTAAATACAAAAGATTTTAAATGCTTTTATTCGCCCTTAGGGGACGGA 4656
DB	5572	TGACTATATTTTGGTAAATACAAAAGATTTTAAATGCTTTTATTCGCCCTTAGGGGACGGA 5513
QY	4657	GTTTGGAAATTTTAAATACGAATGATGTAAGAGATTTGTTTCAACAAACTCTAGCGCTT 4716
DB	5512	GTTTGGAAATTTTAAATACGAATGATGTAAGAGATTTGTTTCAACAAACTCTAGCGCTT 5453
QY	4717	TAGATTTAGTTTGGGGCAATTTATCAGTGTGATATAGCAAGTTTATGCAAAAGCTTCT 4776
DB	5452	TAGATTTAGTTTGGGGCAATTTATCAGTGTGATATAGCAAGTTTATGCAAAAGCTTCT 5393
QY	4777	TGATGCTGAAACGGGGGAAATTTAAATGACAAAACAAAAGAAAAAGAGTTATTCTGAAAAAT 4836
DB	5392	TGATGCTGAAACGGGGGAAATTTAAATGACAAAACAAAAGAAAAAGAGTTATTCTGAAAAAT 5333
QY	4837	GAGGAATTTAAAAAAGAAATTTAAGGCTTAAAGAGCGTTATTCGAAGATACAGAGAAATG 4896
DB	5332	GAGGAATTTAAAAAAGAAATTTAAGGCTTAAAGAGCGTTATTCGAAGATACAGAGAAATG 5273



Qy 4897 GAAGTTGAATTAAGTCAACAATAGATTATTAGAGAGGAGGATATTGATAATAATAAAA 4956  
 Db 5272 GAAGTTGAATTAAGTCAACAATAGATTATTAGAGAGGAGGATATTGATAATAATAAAA 5213

Qy 4957 GCCCCTGACGAAAGTCGGCACTTCGTTCTTTT 4992  
 Db 5212 G-CCCCCTGACGAAAGTCGAAAGGGGGTTCCTTTT 5178

RESULT 8  
 AF076213/c 8142 bp DNA linear SYN 08-FEB-2000  
 LOCUS AF076213 Promoter screenings vector pMM225, complete sequence.  
 DEFINITION AF076213 Promoter screenings vector pMM225, complete sequence.  
 ACCESSION AF076213  
 VERSION AF076213.1 GI:5668895  
 SOURCE  
 ORGANISM  
 Promoter screenings vector pMM225  
 Promoter screenings vector pMM225  
 artificial sequences; vectors.  
 REFERENCE  
 1 (bases 1 to 8142)  
 Vriesema,A.J., Brinkman,R., Kok,J., Dankert,J. and Zaat,S.A.  
 Broad-host-range shuttle vectors for screening of regulated  
 promoter activity in viridans group streptococci: isolation of a  
 pH-regulated promoter  
 Appl. Environ. Microbiol. 66 (2), 535-542 (2000)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 20120502  
 10653715  
 REFERENCE  
 2 (bases 1 to 8142)  
 Vriesema,A.J.M., Brinkman,R., Kok,J., Dankert,J. and Zaat,S.A.J.  
 Direct Submission  
 Submitted (02-JUL-1998) Medical Microbiology, Academic Medical  
 Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105  
 AZ, The Netherlands  
 FEATURES  
 source  
 1. 8142  
 /organism="Promoter screenings vector pMM225"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:79020"  
 26..68  
 189..198  
 212..1006  
 /gene="aph-III"  
 /note="promoterless"  
 212..1006  
 /gene="aph-III"  
 /note="APH-III"  
 /codon\_start=1  
 /transl\_table=11  
 /product="3', 5'-aminoglycoside phosphotransferase type  
 III"  
 /protein\_id="AAD46056.1"  
 /db\_xref="GI:5668896"  
 /translation="MAKMRISPELKKLIEKRCVKQTEGMSPAKVVKLVGENENLYLK  
 MTPSRVKTGYDYVEREKDMMLEGLPVKVLHPRHGWGNLLMSADGVLCEEY  
 EDQSPKGTIELVACIRLPHSIDSCPYTNSLSRLAELDYLLNLDADVDCENWE  
 EDTPDKPRLYDFLTKPEPEELVFSGLDGSNIFVKDQKVSFIDLGRGRADKW  
 YDIAFCVRSIREDBEQVVELFDDLLGKPDWEKIKYILLDELF"  
 1030..4520  
 /note="trp-lacZ fusion"  
 1254..1257  
 1265..4339  
 /gene="lacZ"  
 /note="promoterless"  
 1265..4339  
 /gene="lacZ"  
 /codon\_start=1  
 /transl\_table=11  
 /product="beta-d-galactosidase"  
 /protein\_id="AAD46057.1"  
 /db\_xref="GI:5668897"  
 /translation="MTNITDSLAVIQQRDWNPNQVGTQNLRLAAHPFPFAWRNSBEAR  
 TDRPSQQLRSLNGEWFAPFPAPPEVFSLEWLECDLPEDATVVVPSNQMHGYDAPIYT

Qy 4897 GAAGTTGAATTAAGTCAACAATAGATTATTAGAGAGGAGGATATTGATAATAATAAAA 4956  
 Db 5272 GAAGTTGAATTAAGTCAACAATAGATTATTAGAGAGGAGGATATTGATAATAATAAAA 5213

Qy 4957 GCCCCTGACGAAAGTCGGCACTTCGTTCTTTT 4992  
 Db 5212 G-CCCCCTGACGAAAGTCGAAAGGGGGTTCCTTTT 5178

RESULT 8  
 AF076213/c 8142 bp DNA linear SYN 08-FEB-2000  
 LOCUS AF076213 Promoter screenings vector pMM225, complete sequence.  
 DEFINITION AF076213 Promoter screenings vector pMM225, complete sequence.  
 ACCESSION AF076213  
 VERSION AF076213.1 GI:5668895  
 SOURCE  
 ORGANISM  
 Promoter screenings vector pMM225  
 Promoter screenings vector pMM225  
 artificial sequences; vectors.  
 REFERENCE  
 1 (bases 1 to 8142)  
 Vriesema,A.J., Brinkman,R., Kok,J., Dankert,J. and Zaat,S.A.  
 Broad-host-range shuttle vectors for screening of regulated  
 promoter activity in viridans group streptococci: isolation of a  
 pH-regulated promoter  
 Appl. Environ. Microbiol. 66 (2), 535-542 (2000)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 20120502  
 10653715  
 REFERENCE  
 2 (bases 1 to 8142)  
 Vriesema,A.J.M., Brinkman,R., Kok,J., Dankert,J. and Zaat,S.A.J.  
 Direct Submission  
 Submitted (02-JUL-1998) Medical Microbiology, Academic Medical  
 Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105  
 AZ, The Netherlands  
 FEATURES  
 source  
 1. 8142  
 /organism="Promoter screenings vector pMM225"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:79020"  
 26..68  
 189..198  
 212..1006  
 /gene="aph-III"  
 /note="promoterless"  
 212..1006  
 /gene="aph-III"  
 /note="APH-III"  
 /codon\_start=1  
 /transl\_table=11  
 /product="3', 5'-aminoglycoside phosphotransferase type  
 III"  
 /protein\_id="AAD46056.1"  
 /db\_xref="GI:5668896"  
 /translation="MAKMRISPELKKLIEKRCVKQTEGMSPAKVVKLVGENENLYLK  
 MTPSRVKTGYDYVEREKDMMLEGLPVKVLHPRHGWGNLLMSADGVLCEEY  
 EDQSPKGTIELVACIRLPHSIDSCPYTNSLSRLAELDYLLNLDADVDCENWE  
 EDTPDKPRLYDFLTKPEPEELVFSGLDGSNIFVKDQKVSFIDLGRGRADKW  
 YDIAFCVRSIREDBEQVVELFDDLLGKPDWEKIKYILLDELF"  
 1030..4520  
 /note="trp-lacZ fusion"  
 1254..1257  
 1265..4339  
 /gene="lacZ"  
 /note="promoterless"  
 1265..4339  
 /gene="lacZ"  
 /codon\_start=1  
 /transl\_table=11  
 /product="beta-d-galactosidase"  
 /protein\_id="AAD46057.1"  
 /db\_xref="GI:5668897"  
 /translation="MTNITDSLAVIQQRDWNPNQVGTQNLRLAAHPFPFAWRNSBEAR  
 TDRPSQQLRSLNGEWFAPFPAPPEVFSLEWLECDLPEDATVVVPSNQMHGYDAPIYT

NVTYPLTVNPPFVPTENPTGCSYLTFNVDSEMLQEGQTRIIPDVNSAFLWCNGRW  
 GYQDSRLPSEFDLSAFLRAGENLAVMLRYMSDSYLEDQDMWRMSGIIPDVYSLAHK  
 PTCISDFHATRFNDDFSAVLEAEVQMCGLRDYLRVTYSLWQETQVAGSTAFSG  
 GEIIDRGGYADRVTLRLNVENPKLWSAEIPNLRYRAVVELHTADGLTILAEACDVGFR  
 EVRIENGLLNGKPLLRGVNRHEHPLHGLQVMDQETWQDILLMKQNNFNAVCUSH  
 YPNHMYTLCDRYGLYVVDDEANIETHGMVPMNRLTDDPRMLFAMSERVTRAVQDRN  
 HPSVIWISLNGESHGANDHLYRWIKSVDPSPVOYEGGADTTATDIIICPMYARDV  
 EDOFPAPKWSIKKMLSPGETRPLILCEYAHMGNLSLGGFAKYQWAFROYPRLOCG  
 FVMDVDSLIKXDENGNSPKSAYGGDEGTDENDROFCMGLAFADPTLPAPOGKOLIE  
 QQFQRLSQITIEVSEYLFASDNEHLLHMYVALDGKPLASGEVPLDVAPOGKOLIE  
 LPFLPQESAGQLWTLRVVQPNATAWSEAGHISAWQWRLAENLSVTLPASHALPH  
 LTTSEMDFCLELGNKQNFNRQSGFLSQMWIGDKQLLTPLRDQFTRAPLDNDIGVSE  
 ARIDPNMAVERWKAACHYQAEALLOCTADTLADAVLITTAHAWHQSKTLFISRT  
 YRIDGQMAITVDVEVASDTPHPARIGLNCQLAQVAERVMGLGQSNYDRLTAA  
 CDFRWDLPLSDMYTPVYFSPSENGLCRGTRELNYGPHQWRGDFQFNISYSQQOLMETS  
 HRHLLAEECTWLNDGFMHIGIGDDSDSPSVSAELQLSAGRYHYQLVWCQK  
 complement (5425..6123)  
 /gene="repA"  
 complement (5425..6123)  
 /gene="repA"  
 /codon\_start=1  
 /transl\_table=11  
 /product="replication protein RepA"  
 /protein\_id="AAD46058.1"  
 /db\_xref="GI:5668898"  
 /translation="MAIKNTKARNFQFLYPDSIPNDWKEKLESGLSVMAVSLPHMD  
 EKDKOTWNSDVIRNGHYKPHYVIVIAENPVITIEYRNKIKKLGNSVLAHYEI  
 LDYIKGSYEYLTHTESDAKAKKHIDYKDIILNDFIDRIDYTLDESQKRELKLL  
 DIVDDYLVNTKDLMAFIRLGAEGFGLINTNDVKDIVSTNSSAFRLWFEFGYCCGVA  
 SYAKVLDARTGEIK"  
 7216..7275  
 /codon\_start=1  
 /transl\_table=11  
 /product="emrC leader peptide"  
 /protein\_id="AAD46059.1"  
 /db\_xref="GI:5668899"  
 /translation="MGIFSIFFVISTVHYQPNKK"  
 7335..8069  
 /gene="emrC"  
 7335..8069  
 /gene="emrC"  
 /codon\_start=1  
 /transl\_table=11  
 /product="adenine methylase"  
 /protein\_id="AAD46060.1"  
 /db\_xref="GI:5668900"  
 /translation="MNEKNIKHSQNFITSKHNDIKMTNIRLNEHNIPEIGSKGHP  
 TLELVKRCNFVTAIEIDHKLCKTENKLVHDNDFOVLNKDIMQFKPKQSKYKYN  
 PYNISTDIIRKIVFDSIANEIVLIVEYGFARLLNTRSLALLMAEVIDISILSMVPR  
 EYFHPKPNVSSLIRLSRKSRIKSHDKQKQYVFMKWNVNKVKKIFTKNQFNLSLKH  
 AGIDDLNLSIFEQFLSFLNSYKLFNK"

ORIGIN  
 Query Match 31.98; Score 1641.6; DB 12; Length 8142;  
 Best Local Similarity 98.48; Pred. No. 1e-263;  
 Matches 1689; Conservative 0; Mismatches 24; Indels 3; Gaps 3;  
 Qy 3279 TTTATATTTATTTTCGGTATTTTATTAACAGCTGCTCAAAATCGTTCTCGGAGCGTTT 3338  
 Db 6952 TATATATTTATTTATTCGGATTTTATTAACAGCTGCTCAAAATCGTTCTCGGAGCGTTT 6993  
 Qy 3339 TAGCGTTTATTTTCGTTAGTTATCGGCATATTCGTTAAAAACAGCGGTTATCGTAGCGTAA 3398  
 Db 6992 TAGCGTTTATTTTCGTTAGTTATCGGCATATTCGTTAAAAACAGCGGTTATCGTAGCGTAA 6933  
 Qy 3399 AAGCCCTTGACGTAGCGT-GCTTTCGACGAGATGCTTCTGTTAGATTATGAAGCC 3457  
 Db 6832 AAGCCCTTGACGTAGCGTGGCTTTGACGCGAAGATGTTGCTGTTAGATATGAAGCC 6773  
 Qy 3458 GATGACTGAATGAATAATAAGCGCAGCGTCCCTTCTATTTTCGGTTGAGGAGGCTCAAGG 3517  
 Db 6772 GATGACTGAATGAATAATAAGCGCAGCGCCCTTCTATTTTCGGTTGAGGAGGCTCAAGG 6713

QY 3518 GAGTTTGAGGGATGAATTCCTCATGGGTTTGATTTTAAATAATGCTTGCAATTTTGC 3577  
DB 6712 GAGTATGAGGGAGTGAATTCCTCATGGGTTTGATTTTAAATAATGCTTGCAATTTTGC 6653  
QY 3578 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTGGAATTTGGAATAAATGAGGGGA 3637  
DB 6652 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTGGAATTTGGAATAAATGAGGGGA 6593  
QY 3638 AAGGAACGAATTTTCTTCCGTACTACGACCCCAATTAAGTCGCGAGTGCCTAATTTT 3697  
DB 6592 AAGGAACGAATTTTCTTCCGTACTACGACCCCAATTAAGTCGCGAGTGCCTAATTTT 6533  
QY 3698 GTCCCAAAAACGCTCTATCCCACTCGCTCAAGGGTTTGAAGGGTTTCAATTCGCCAAC 3757  
DB 6532 GTCCCAAAAACGCTCTATCCCACTCGCTCAAGGGTTTGAAGGGTTTCAATTCGCCAAC 6473  
QY 3758 GAATCCCAACGTTTTCGCCAAGCTTTTATTAATCTATATTAAGTAGCTTATTTGTT 3817  
DB 6472 GAATCCCAACGTTTTCGCCAAGCTTTTATTAATCTATATTAAGTAGCTTATTTGTT 6413  
QY 3818 GTTTTATGATTAACAAAGTACACTAAATTTTATAAATTTATTTGATTTGAGTTTATA 3877  
DB 6412 GTTTTATGATTAACAAAGTACACTAAATTTTATAAATTTATTTGATTTGAGTTTATA 6353  
QY 3878 AATGGTGAATTCAGAAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAAGACAGATAA 3937  
DB 6352 AATGGTGAATTCAGAAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAAGACAGATAA 6293  
QY 3938 ARAATTAACAGATATCGGAAACAAAAAGGTTTTCAAAATCTGCGGTTGCGCGTAGC 3997  
DB 6292 ARAATTAACAGATATCGGAAACAAAAAGGTTTTCAAAATCTGCGGTTGCGCGTAGC 6233  
QY 3998 TATAGAAGATATCGAAGAAAGGAATCAGAACAAAAAATAAGACGAAGCTCGCGTTT 4057  
DB 6232 TATAGAAGATATCGAAGAAAGGAATCAGAACAAAAAATAAGACGAAGCTCGCGTTT 6173  
QY 4058 TAGAAGATACGAGTTTTCCTACTTCTTTTATAGTAAATTAATATATGCTCTATTA 4116  
DB 6172 TAGAAGATACGAGTTTTCCTACTTCTTTTATAGTAAATTAATATATGCTCTATTA 6113  
QY 4117 AATACTAAAGCTAGAAATTTTGATTTTATATATCTGACTCAATTCCTAATGATTG 4176  
DB 6112 AATACTAAAGCTAGAAATTTTGATTTTATATATCTGACTCAATTCCTAATGATTG 6053  
QY 4177 GAAAGAAAAATAGAGATTTGGCGGTATCTATGCGCTGACGCTCTTACACATATGA 4236  
DB 6052 GAAAGAAAAATAGAGATTTGGCGGTATCTATGCGCTGACGCTCTTACACATATGA 5993  
QY 4237 CGAAAAAAGATAAAGATACATCGAATAGTAGTATGTTATACGAATCGAAAGCACTA 4296  
DB 5992 CGAAAAAAGATAAAGATACATCGAATAGTAGTATGTTATACGAATCGAAAGCACTA 5933  
QY 4297 TAAAAAACCACTATACGTTATATATATTCACGAAATCCTGTAAACAATAGAACGCT 4356  
DB 5932 TAAAAAACCACTATACGTTATATATATTCACGAAATCCTGTAAACAATAGAACGCT 5873  
QY 4357 TAGGAACAGATTAAGGAAATTTGGGAATAGTTCATGCTCATCTTCAGATATCTGA 4416  
DB 5872 TAGGAACAGATTAAGGAAATTTGGGAATAGTTCATGCTCATCTTCAGATATCTGA 5813  
QY 4417 TTATATCAAGGTTTCATATGAATTTTGACTCATGAATCAAGGACGCTATTGCTAGAA 4476  
DB 5812 TTATATCAAGGTTTCATATGAATTTTGACTCATGAATCAAGGACGCTATTGCTAGAA 5753  
QY 4477 TAAACATATACGCAAAAAAGATATTTTGAACATTAATGATTTTGATATGACCGCTA 4536  
DB 5752 TAAACATATACGCAAAAAAGATATTTTGAACATTAATGATTTTGATATGACCGCTA 5693  
QY 4537 TATACACTTCATGAAGCCAAAAAGAGATTTGAAGATTTTACTTTTAGATATAGTGA 4596  
DB 5692 TATACACTTCATGAAGCCAAAAAGAGATTTGAAGATTTTACTTTTAGATATAGTGA 5633  
QY 4597 TGACTATAATTTGGTAAATACAAAAGATTTAATGCTTTTATTCGCTTAGGGAGCGGA 4656

DB 5632 TGACTATAATTTGGTAAATACAAAAGATTTAATGCGCTTTTATTGCGCTTAGGGAGCGGA 5573  
QY 4657 GTTTGGAAATTTAAATACGAATGATGATAAAGATATTGTTTCAACAACTCTAGCGCTT 4716  
DB 5572 GTTTGGAAATTTAAATACGAATGATGATAAAGATATTGTTTCAACAACTCTAGCGCTT 5513  
QY 4717 TAGATTATGTTTGAAGGCAATTAATCAGTGTGGATATAGAGCAAGTTATGCAAAAGTTCT 4776  
DB 5512 TAGATTATGTTTGAAGGCAATTAATCAGTGTGGATATAGAGCAAGTTATGCAAAAGTTCT 5453  
QY 4777 TGATCTGTAACGCGGGGAATAAATCACAACAAAGAAAGAGTTATTGCTGAAAAAT 4836  
DB 5452 TGATCTGTAACGCGGGGAATAAATCACAACAAAGAAAGAGTTATTGCTGAAAAAT 5393  
QY 4837 CAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 4896  
DB 5392 CAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 5333  
QY 4897 GAAGTTGAATTAAGTACAACTAGATTTATTGAGAGGAGGATTTGTAATAATAAAA 4956  
DB 5332 GAAGTTGAATTAAGTACAACTAGATTTATTGAGAGGAGGATTTGTAATAATAAAA 5273  
QY 4957 GCGCCCTCGACGAAAGTCGCGACTTCGTTCTTTTT 4992  
DB 5272 G-CGCTGACGAAAGTCGAGGGGTTTTTATTTT 5238

RESULT 9  
LLREPA L.lactis plasmid pw01 2178 bp DNA linear BCT 18-OCT-1991  
LOCUS ORFs B, C, and D.  
DEFINITION  
ACCESSION X56954.1 GI:44058  
VERSION repA gene; replication protein; replication protein A.  
KEYWORDS Lactococcus lactis  
SOURCE Lactococcus lactis  
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

REFERENCE 1  
AUTHORS Leenhouts,K.J., Tolner,B., Bron,S., Kok,J., Venema,G. and Seegers,J.F.  
TITLE Nucleotide sequence and characterization of the broad-host-range lactococcal plasmid pw01  
JOURNAL Plasmid 26 (1), 55-66 (1991)  
MEDLINE 92052683  
PUBMED 1840693  
REFERENCE 2 (bases 1 to 2178)  
AUTHORS Leenhouts,K.J.  
TITLE Direct Submission  
JOURNAL Submitted (10-DEC-1990) K.J. Leenhouts, INSTITUTE OF GENETICS, UNIVERSITY OF GRONINGEN, KERKLAAN 30, 9751 NN HAREN, THE NETHERLANDS

FEATURES  
Source  
1..2178  
/organism="Lactococcus lactis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1358"  
/plasmid="pw01"  
355..544  
/gene="ORF B"  
335..544  
/gene="ORF B"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA40270.1"  
/db\_xref="GI:44059"  
/db\_xref="GOA:Q48693"  
/db\_xref="SPTREMBL:Q48693"  
/translation="MGKEANFASVLRPIKCRVIFVPKLTLYPNWLKLRGFSANE SPTFSPTFINLYLSFSFVFMITK"  
522..550  
promoter

```

gene      585..746
/genes="ORF C"
CDS       585..746
/genes="ORF C"
/codon_start=1
/transl_table=11
/protein_id="CAA40271.1"
/db_xref="GI:44060"
/db_xref="GOA:Q04136"
/db_xref="SPTREMBL:Q04136"
/translacion="WVISESKKVMISLTKEDQKLTDMAKQKQFSKSAVALAIEEY
ARKESEQKK"
813..1511
/genes="RePA"
813..1511
/genes="RePA"
/codon_start=1
/transl_table=11
/product="replication protein A"
/protein_id="CAA40272.1"
/db_xref="GI:44061"
/db_xref="GOA:Q04138"
/db_xref="SPTREMBL:Q04138"
/translacion="MAIKNTKARNFGFLYPPDISPNWKEKLESVMVSPPLHMD
EKDKDTWSSDVIRNGHKYKXPHVHVIYIENPVTIESVRNKIKRKLNSVAHVEI
LDYKSGEYLTHESKDAKAKKHIIYDKDILINDPDIRYITIDESQKRELKXULL
DIVDDNLVNTKDLMAFIRLGAEPGILNTDKDIVNSSAFRLWFEQNYQCGYRA
SYAKVLDAETGEIK"
1508..1657
/genes="ORF D"
1508..1657
/genes="ORF D"
/codon_start=1
/transl_table=11
/protein_id="CAA40273.1"
/db_xref="GI:44062"
/db_xref="GOA:Q48695"
/db_xref="SPTREMBL:Q48695"
/translacion="MTNKEKELFAENEELKKIKDKRIERYREWEVELSTIIDLLR
GGIIE"
1656..1697
terminator
ORIGIN

Query Match      31.8%; Score 1637; DB 1; Length 2178;
Best Local Similarity 98.9%; Pred. No. 7, 1e-263;
Matches 1679; Conservative 0; Mismatches 15; Indels 3; Gaps 3;

QY 3298 ATTTTATTAACGCTCAAAATCGTTTCGGACGTTTACGCTTTATTCGTTTAC 3357
DB 3 ATTTTATTAACGCTCAAAATCGTTTCGGACGTTTACGCTTTATTCGTTTAC 62
QY 3358 TTATCGCATAAATCGTTTAAACAGCGGTTATCGTAGCTTAAAGCCCTTGACGCTAGCGT 3417
DB 63 TTATCGCATAAATCGTTTAAACAGCGGTTATCGTAGCTTAAAGCCCTTGACGCTAGCGT 122
QY 3418 -GTTTTCACGGAAGATGTGCTGTAGATTATGAAGCCGATGACTGAATGAATAAT 3476
DB 123 GGCTTTCACGGAAGATGTGCTGTAGATTATGAAGCCGATGACTGAATGAATAAT 182
QY 3477 AAGCGAGCGCTCTCTATTCGTTTGGAGGAGCGCTCAAGGGAGTTTGAGGGAATGAAT 3536
DB 183 AAGCGAGCGCGCTCTCTATTCGTTTGGAGGAGCGCTCAAGGGAGTTTGAGGGAATGAAT 242
QY 3537 TCCTCATGCGTTTGAATTTTAAATAATGCTTGAATTTTCCGACGCGTAGCGTGGAAA 3596
DB 243 TCCTCATGCGTTTGAATTTTAAATAATGCTTGAATTTTCCGACGCGTAGCGTGGAAA 302
QY 3597 ATTTTGAATAAATTTGGAATTTGGAATAAATGCGGGAAGGAGCGAATTTTGCTT 3656
DB 303 ATTTTGAATAAATTTGGAATTTGGAATAAATGCGGGAAGGAGCGAATTTTGCTT 362
QY 3657 CCGTACTACGACCCCAATTAAGTGCAGGTGCCAATTTTGTGCAAAACGCTCTATC 3716

```

```

DB 363 CCGTACTACGACCCCAATTAAGTGCAGGTGCCAATTTTGTGCCAAAACGCTCTATC 422
QY 3717 CCAACTGGCTCAAGGGTTTGAGGGGTTTTCAATCGCCAACGAATCGCCAACGTTTTCGC 3776
DB 423 CCAACTGGCTCAAGGGTTTGAGGGGTTTTCAATCGCCAACGAATCGCCAACGTTTTCGC 482
QY 3777 CAACGTTTTTATAAATCTATAATTAAGTAGCTTTATGTGTGTTTTATGATTACAAAGT 3836
DB 483 CAACGTTTTTATAAATCTATAATTAAGTAGCTTTATGTGTGTTTTATGATTACAAAGT 542
QY 3837 GATACACTAAATTTTATAAATTAATTCATTGGAGTTTTTAAATGGTGTATTCAGAAATCG 3896
DB 543 GATACACTAAATTTTATAAATTAATTCATTGGAGTTTTTAAATGGTGTATTCAGAAATCG 602
QY 3897 AAAAAAAGAGTTATGATTTCTCTGACAAAAAGACAGATAAAAAATTAACAGATATGGCG 3956
DB 603 AAAAAAAGAGTTATGATTTCTCTGACAAAAAGACAGATAAAAAATTAACAGATATGGCG 662
QY 3957 AACAAAAAGGTTTTTCAAAATCTCGGTTGCGGCTTAGCTATAGAAGATATGCAAGA 4016
DB 663 AACAAAAAGGTTTTTCAAAATCTCGGTTGCGGCTTAGCTATAGAAGATATGCAAGA 722
QY 4017 AAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTTTTGAAGGATACGAGTTTC 4076
DB 723 AAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTTTTGAAGGATACGAGTTTC 782
QY 4077 GCTACTTGTGTTTGAATAGGTAA-TATATCATGCTATTAATAAATACTAAAGCTAGAAAT 4135
DB 783 GCTACTTGTGTTTGAATAGGTAAATATATCATGCTATTAATAAATACTAAAGCTAGAAAT 842
QY 4136 TTTGGAATTTTATTAATCTCTGACTCAATTCCTAAATGCTTGAAGAAAAAATTAGAGAGT 4195
DB 843 TTTGGAATTTTATTAATCTCTGACTCAATTCCTAAATGCTTGAAGAAAAAATTAGAGAGT 902
QY 4196 TTGGGCTATCTATGGCTGTCAGTCTTTACAGATATGGACGAAAAAAGATTAAGAT 4255
DB 903 TTGGGCTATCTATGGCTGTCAGTCTTTACAGATATGGACGAAAAAAGATTAAGAT 962
QY 4256 ACATGGAATAGTAGTGTATTAACGAAATGGAAGACACTATAAAAACACACTATCAC 4315
DB 963 ACATGGAATAGTAGTGTATTAACGAAATGGAAGACACTATAAAAACACACTATCAC 1022
QY 4316 GTTATATATATTCACGAAATCTGTAAACAAATAGAACGCTTAGAACCAAGATTAAGCGA 4375
DB 1023 GTTATATATATTCACGAAATCTGTAAACAAATAGAACGCTTAGAACCAAGATTAAGCGA 1082
QY 4376 AAATTTGGGGAATAGTTCAGTGTGCTCATGTTGAGTACTTGTATATATCAAGGTTTCATAT 4435
DB 1083 AAATTTGGGGAATAGTTCAGTGTGCTCATGTTGAGTACTTGTATATATCAAGGTTTCATAT 1142
QY 4436 GAATATTTGACTCATGAATCAAAAGGACGCTATTGCTAAGAAATAAACATATATACGCAAA 4495
DB 1143 GAATATTTGACTCATGAATCAAAAGGACGCTATTGCTAAGAAATAAACATATATACGCAAA 1202
QY 4496 AAAGATATTTGACATTAATGATTTTGTATTTGACCGCTATATAACACTTGTATGAAGC 4555
DB 1203 AAAGATATTTGACATTAATGATTTTGTATTTGACCGCTATATAACACTTGTATGAAGC 1262
QY 4556 CAAAAAGAGAAATTAAGAAATTTTACTTTTAGATATATAGTGGATGACTATAATTTTGTAAAT 4615
DB 1263 CAAAAAGAGAAATTAAGAAATTTTACTTTTAGATATATAGTGGATGACTATAATTTTGTAAAT 1322
QY 4616 AAAAAAGATTTAATGGCTTTTATTCGCTTAGGGGACGAGTTTGGATTTTAAATAGC 4675
DB 1323 AAAAAAGATTTAATGGCTTTTATTCGCTTAGGGGACGAGTTTGGATTTTAAATAGC 1382
QY 4676 AATGATGTAAGAGATATTTGTTTCAAAACACTCTAGCGCTTTTAGATTATGTTTGGAGGC 4735
DB 1383 AATGATGTAAGAGATATTTGTTTCAAAACACTCTAGCGCTTTTAGATTATGTTTGGAGGC 1442
QY 4736 AATTATCAGTGTGGATATAGAGCAAGTTTATGCAAGGTTTCTTGATCTGCTGAACCGGGGAA 4795
DB 1443 AATTATCAGTGTGGATATAGAGCAAGTTTATGCAAGGTTTCTTGATCTGCTGAACCGGGGAA 1502

```

```
Qy 4796 ATAAATGACAAACAAGAAAAAGATTATTTCTGTAATAGAGGAATTAATAAAGAAA 4855
Db 1503 ATAAATGACAAACAAGAAAAAGATTATTTCTGTAATAGAGGAATTAATAAAGAAA 1562
Qy 4856 TTAAGGACTTAAAGAGCGTATTGAAGATACAGAGAAATGGAAGTTGAATTAAGTACAA 4915
Db 1563 TTAAGGACTTAAAGAGCGTATTGAAGATACAGAGAAATGGAAGTTGAATTAAGTACAA 1622
Qy 4916 CAATAGATTATTGAGAGGAGGATTATTGAATTAATAAAGCCCTTGACGAAAGTCG 4975
Db 1623 CAATAGATTATTGAGAGGAGGATTATTGAATTAATAAAG-CCCTTGACGAAAGTCG 1681
Qy 4976 CGACTTCGTTCTTTT 4992
Db 1682 AAGGGGTTTATTATTT 1698

RESULT 10
AF372620 5126 bp DNA circular SYN 31-MAY-2002
LOCUS GlnQ allelic exchange vector pAG101, complete sequence.
ACCESSION AF372620
VERSION AF372620.1 GI:14211230
KEYWORDS
SOURCE glnQ allelic exchange vector pAG101
ORGANISM glnQ allelic exchange vector pAG101
REFERENCE 1 (bases 1 to 5126)
AUTHORS Tamura,G.S., Nittayajarn,A. and Schoentag,D.L.
TITLE A glutamine transport gene, glnQ, is required for fibronectin
adherence and virulence of group B streptococci
JOURNAL Infect. Immun. 70 (6), 2877-2885 (2002)
MEDLINE 22006888
PUBMED 12010975
REFERENCE 2 (bases 1 to 5126)
AUTHORS Tamura,G.S., Schoentag,D.L. and Nittayajarn,A.
TITLE Direct Submission
JOURNAL Submitted (20-APR-2001) Department of Pediatrics, Division of
Infectious Diseases, Children's Hospital and Regional Medical
Center and the University of Washington, 4800 Sand Point Way NE,
Seattle, WA 98105-0371, USA
FEATURES
source
1. .5126
/organism="glnQ allelic exchange vector pAG101"
/mol_type="genomic DNA"
/db_xref="taxon:159742"
/lab_host="Streptococcus agalactiae"
/plasmid="pAG101"
/notes="temperature-sensitive plasmid"
order(4747..5126,859..1116)
/notes="glnQ gene interrupted by erm gene"
102..839
/gene="erm"
102..839
/gene="erm"
/notes="confers erythromycin-resistance"
/codon_start=1
/transl_table=11
/product="ribosomal RNA methylase"
/protein_id="AAK57385.1"
/db_xref="GI:14211231"
/translation="MNFNKIDLDNWKRKEIPNHYLNQOTFSTITEIDISVLVYNNIKQ
TKLAKISQVTSIBLSDSHLSEKULKNTVLIHQDILQFQFNKQRYKIVGSIP
YHLSTQIKVVFESHADYLIIVEEGFYKRLDIDHRTGLLLHTQVSIQOLKJFAB
CFHPKPKVNSYLIILKTRHTDVPDKYWKLYTFVSKWVNREYRQLFTKNQFHQAMKHA
KVNLSVTYVQVLSIFNSYLLFNGRK"
3845..4495
/gene="cat"
3845..4495
/gene="cat"
/notes="confers chloramphenicol resistance"
/codon_start=1
```

```
/transl_table=11
/product="chloramphenicol acetyl transferase"
/protein_id="AAK57386.1"
/db_xref="GI:14211232"
/translation="MNFNKIDLDNWKRKEIPNHYLNQOTFSTITEIDISVLVYNNIKQ
EGYFYFAFIFLVRVNTSNTAFRTYNSDGLGYWDLKLELYTIFPGVSTFGSIWT
PVKNDKFEFDLYLSDVEKNGSKLPKTPIPENAFSLSIIPWTSFTGFLNINNS
NYLLPIITAGNFINKNSIYLPLSLQVHSHVCDGVHAGLFMNSIQELSDRNDWLL"
```

## ORIGIN

```
Query Match 31.8%; Score 1635.2; DB 12; Length 5126;
Best Local Similarity 98.2%; Pred. No. 1.3e-262;
Matches 1685; Conservative 0; Mismatches 28; Indels 3; Gaps 3;

Qy 3279 TTTATATTTATTTTCGGTATTTTATTATAAAGCTCTCAAAATCGTTCTCTGGACGTTT 3338
Db 1395 TATATATATTTATATATCCGATTTTATTATAAAGCTCTCAAAATCGTTCTGGACGTTT 1454
Qy 3339 TAGCGTTTATTTTCGGTTAGTTATCGGCATAATCGTTATAAAGCGGCTTATCGTAGCGTAA 3398
Db 1455 TAGCGTTTATTTTCGGTTAGTTATCGGCATAATCGTTATAAAGCGGCTTATCGTAGCGTAA 1514
Qy 3399 AAGCCCTTTGAGCGTAGCGT-GCTTTGCAGCGAAGATGTTGCTGTAGATATTAAGAGCC 3457
Db 1515 AAGCCCTTTGAGCGTAGCGTGTTCAGCGAAGATGTTGCTGTAGATATTAAGAGCC 1574
Qy 3458 GATGACTGAATCAATAATAAAGCGCAGCGCTCTCTATTTCGGTTGGAGGCTCAAGG 3517
Db 1575 GATGACTGAATCAATAATAAAGCGCAGCGCTCTATTTCGGTTGGAGGCTCAAGG 1634
Qy 3518 GAGTTGAGGGAATGAATTCCTCATCGGTTTGNATTTTAAATTTGTTTGCATTTTGC 3577
Db 1635 GAGTATGAGGGAATGAATTCCTCATCGGTTTGNATTTTAAATTTGTTTGCATTTTGC 1694
Qy 3578 CGAGCGGTAGCGCTCGAAAAATTTTGAATAAATTTGGAATTTGAAAAAATGCGGGGA 3637
Db 1695 CGAGCGGTAGCGCTCGAAAAATTTTGAATAAATTTGGAATTTGAAAAAATGCGGGGA 1754
Qy 3638 AAGGAAGCGAATTTTTCGCTTCGCTACTAGACCCCTTAAGTCCGAGTGCCAAATTTT 3697
Db 1755 AAGGAAGCGAATTTTTCGCTTCGCTACTAGACCCCTTAAGTCCGAGTGCCAAATTTT 1814
Qy 3698 GTGCCAAAAACGCTCTATCCCACTGGCTCAAGGGTTTGAGGGTTTTCATCCCAAC 3757
Db 1815 GTGCCAAAAACGCTCTATCCCACTGGCTCAAGGGTTTGAGGGTTTTCATCCCAAC 1874
Qy 3758 GAATCGCCAAACGTTTTTCGCCAACGTTTTTATAAATCTATATTTAAGTAGCTTTATGTT 3817
Db 1875 GAATCGCCAAACGTTTTTCGCCAACGTTTTTATAAATCTATATTTAAGTAGCTTTATGTT 1934
Qy 3818 GTTTTATGATTAAGAAGTATACACTAATTTTATAAATTTTATGATTTGAGTTTATA 3877
Db 1935 GTTTTATGATTAAGAAGTATACACTAATTTTATAAATTTTATGATTTGAGTTTATA 1994
Qy 3878 AATGGTGATTTTCAGAAATCGAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 3937
Db 1995 AATGGTGATTTTCAGAAATCGAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 2054
Qy 3938 AAAATTAACAGATATGGCGAAACAAAAAGTTTTTCAAAATCTCGGTTGCGCGCTTAGC 3997
Db 2055 AAAATTAACAGATATGGCGAAACAAAAAGTTTTTCAAAATCTCGGTTGCGCGCTTAGC 2114
Qy 3998 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTT 4057
Db 2115 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTT 2174
Qy 4058 TAGAAGGATACGAGTTTTCGCTACTGTTGTTTGAAGTAA-TATATCATCGGCTATTAA 4116
Db 2175 TAGAAGGATACGAGTTTTCGCTACTGTTGTTTGAAGTAA-TATATCATCGGCTATTAA 2234
Qy 4117 AAATACTAAGCTAGAAATTTTGGATTTTATATATCTGACTCAATCTTAAGTATG 4176
Db 2235 AAATACTAAGCTAGAAATTTTGGATTTTATATATCTGACTCAATCTTAAGTATG 2294
```

QY	4177	GAAGAAATTTAGAGAGTTTGGCGTATCTATGCTGTCACTCTCTTTACAGATATGGA	4236	JOURNAL	Appl. Environ. Microbiol. 63 (9), 3539-3547 (1997)
Db	2295	GAAGAAATTTAGAGAGTTTGGCGTATCTATGCTGTCACTCTCTTTACAGATATGGA	2354	MEDLINE	97438516
QY	4237	CGAAAAAAGATAGATACATGATAGTGTATACGAATGGAACACTA	4296	PUBMED	9293004
Db	2355	CGAAAAAAGATAGATACATGATAGTGTATACGAATGGAACACTA	2414	REFERENCE	2 (bases 1 to 9021)
QY	4297	TAAAAACACACATATCATGTTATATATTCACCAATCTCTGTAACAATAGAAAGCGT	4356	AUTHORS	Jones, A.L. and Rubens, C.E.
Db	2415	TAAAAACACACATATCATGTTATATATTCACCAATCTCTGTAACAATAGAAAGCGT	2474	TITLE	Direct Submission
QY	4357	TAGGAACAAGATTAAAGCGAAATTTGGGAATAGTTTCAGTTGCTCATGTTGAGATACCTTGA	4416	JOURNAL	Submitted (23-APR-1998) Pediatrics, Children's Hospital and Regional Medical Center, 4800 Sand Point Way N.E., Seattle, WA 98105, U.S.A.
Db	2475	TAGGAACAAGATTAAAGCGAAATTTGGGAATAGTTTCAGTTGCTCATGTTGAGATACCTTGA	2534	FEATURES	Location/Qualifiers
QY	4417	TTATATCAAAAGTTTCATATGAATATTGACTCATGAATCAAAAGGACGCTATTGCTTAAGAA	4476	source	1. 9021
Db	2535	TTATATCAAAAGTTTCATATGAATATTGACTCATGAATCAAAAGGACGCTATTGCTTAAGAA	2594	source	/organism="Tn917 delivery vector pAJ005"
QY	4477	TAAACATATATACGACAAAAGATATTGTAACATTAATGATTTTGTATTTGACCGCTA	4536	gene	/mol_type="genomic DNA"
Db	2595	TAAACATATATACGACAAAAGATATTGTAACATTAATGATTTTGTATTTGACCGCTA	2654	CDS	/db_xref="taxon:76156"
QY	4537	TATACACTTGTATGAAGCCMAAAAGAGAAATTTAGAGATTTTACTTTTATAGATATAGTGA	4596	gene	/lab_host="Streptococcus agalactiae A909"
Db	2655	TATACACTTGTATGAAGCCMAAAAGAGAAATTTAGAGATTTTACTTTTATAGATATAGTGA	2714	CDS	/plasmid="pAJ005"
QY	4597	TGACTATAATTTGGTAAATACAAAAGATTTAATGGCTTTTATTCGCTTTAGCGGACCGGA	4656	gene	/transposon="Tn917"
Db	2715	TGACTATAATTTGGTAAATACAAAAGATTTAATGGCTTTTATTCGCTTTAGCGGACCGGA	2774	gene	/genes="erm"
QY	4657	GTTCGAAATTTAAATACGATCATGTAAGATATTGTTCAACAACTCTAGCGCCTT	4716	CDS	/notes="inducible"
Db	2775	GTTCGAAATTTAAATACGATCATGTAAGATATTGTTCAACAACTCTAGCGCCTT	2834	gene	/complement(4167..4928)
QY	4717	TAGATATGTTTGGGCGCAATTTATCAGTGTGATATAGACAAAGTTATGCAAAAGTTCT	4776	CDS	/codon_start=1
Db	2835	TAGATATGTTTGGGCGCAATTTATCAGTGTGATATAGACAAAGTTATGCAAAAGTTCT	2894	gene	/transl_table=11
QY	4777	TGATGCTGAAACGGGGAAATAAAATGACAAACAAAGAAAGATTTTCTGCTGAAAT	4836	CDS	/product="aminoglycoside phosphotransferase A3"
Db	2895	TGATGCTGAAACGGGGAAATAAAATGACAAACAAAGAAAGATTTTCTGCTGAAAT	2954	gene	/protein_id="AAC16050.1"
QY	4837	GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAAATG	4896	CDS	/db_xref="GI:3127118"
Db	2955	GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAAATG	3014	gene	/translations="MAKMRISPCLKKIEKRCVKDTEGMSPAKYVKLVGLENLYLK
QY	4897	GAAATTGAAATTAAGTACAAATAGATTTATTTAGAGAGGAGGATTTGAAATATAAA	4956	CDS	MTSRVGTGYDYVERKDMWLEGLVPVKVLFHFDHGSNLLSEADGVLCSBEY
Db	3015	GAAATTGAAATTAAGTACAAATAGATTTATTTAGAGAGGAGGATTTGAAATATAAA	3074	gene	EDSQSPKIELYAEICIRLPHSDISDCPYTNSLDSLAELDYLLNDLADVDCEKNE
QY	4957	GCCCCCTGACGAAGTCGCGACTTCGTTCTTTT	4992	gene	EDTFKDPRELIDFLKTRPEELVFSHGDLGDSNIFVDKGKVSGLDGRSGRADKW
Db	3075	G-CCCCCTGACGAAGTCGAAGGGGGTTTATTTT	3109	CDS	YDIAFCVRSIRDIGEEQYVELFDFLLGKPKDWEKIKYILLDEL
RESULT 11	AF061336/c	9021 bp	DNA	circular SYN 12-MAY-1998	
LOCUS	AF061336				
DEFINITION	Tn917, delivery vector pAJ005, complete sequence.				
ACCESSION	AF061336				
VERSION	AF061336.1	GI:3127116			
KEYWORDS	Tn917 delivery vector pAJ005				
SOURCE	Tn917 delivery vector pAJ005				
ORGANISM	Tn917 delivery vector pAJ005				
REFERENCE	1 (bases 1 to 9021)				
AUTHORS	Framson, P.E., Nittayajarn, A., Merry, J., Youngman, P. and Rubens, C.E.				
TITLE	New genetic techniques for group B streptococci: high-efficiency transformation, maintenance of temperature-sensitive plasmids, and mutagenesis with Tn917				

/db\_xref="GI:3127120"  
/translation="MWISSEKRVMLSLTKEQDKLTDMAKOKGFSKAVAPALAEY  
ARKSEBQKX"

ORIGIN

Query Match	31.8%	Score	1635.2	DB 12	Length	9021
Best Local Similarity	98.2%	Pred. No.	1.2e-262			
Matches 1685	Conservative	0	Mismatches	28	Indels	3
Gaps						3
3279	TTTATATTTATTTTCGGTATTTTATTAATAACGCTCAAAATCGTTTCTGGGAGCTTT	3338				
Db	8804	TATATATTTATTTATTCGGATTTTATTAATAACGCTCAAAATCGTTTCTGGGAGCTTT	8745			
3339	TAGCGTTATTTCTGTTAGTATCGCATATCGTTAAACAGCGCTATCGTAGCGTAA	3398				
Db	8744	TAGCGTTATTTCTGTTAGTATCGCATATCGTTAAACAGCGCTATCGTAGCGTAA	8685			
3399	AAGCCCTTGAGCGTAGCGT-GCTTTGAGCGAAGATGTTGTCTGTGTAGATTATGAAGCC	3457				
Db	8684	AAGCCCTTGAGCGTAGCGTGTGTCAGCGAAGATGTTGTCTGTGTAGATTATGAAGCC	8625			
3458	GATGACTGAATGAATTAATGAAGCGAGCGCTCTCTATTTTCGTTGCGAGGCGCTCAAG	3517				
Db	8624	GATGACTGAATGAATTAATGAAGCGAGCGCTCTCTATTTTCGTTGCGAGGCGCTCAAG	8565			
3518	GAGTTTCAGGGAATGAATTCCTCATGCGTTTGAATTTTAAAAATTCGTTGCAATTTGC	3577				
Db	8564	GAGTTTCAGGGAATGAATTCCTCATGCGTTTGAATTTTAAAAATTCGTTGCAATTTGC	8505			
3578	CGAGCGTAGCGTGGAAAAATTTTGAATAAATTTGGAATTTGGAAAAAATTTGGGGGA	3637				
Db	8504	CGAGCGTAGCGTGGAAAAATTTTGAATAAATTTGGAATTTGGAAAAAATTTGGGGGA	8445			
3638	AAGAAAGCAATTTGCTTCGCTACGACCGCCCTTAAAGTCCGAGTGCCTCAATTTT	3697				
Db	8444	AAGAAAGCAATTTGCTTCGCTACGACCGCCCTTAAAGTCCGAGTGCCTCAATTTT	8385			
3698	GTGCCAAAAACGCTCTATCCCACTGCTCAAGGGTTTGAGGGTTTTCATTCGCCAAC	3757				
Db	8384	GTGCCAAAAACGCTCTATCCCACTGCTCAAGGGTTTGAGGGTTTTCATTCGCCAAC	8325			
3758	GAATCGCAAGCTTTTCGCCAAGCTTTTATATATCTATATTAAGTCTATATGTT	3817				
Db	8324	GAATCGCAAGCTTTTCGCCAAGCTTTTATATATCTATATTAAGTCTATATGTT	8265			
3818	GTTTTATGATTACAAAGTGATACACTAAATTTTATAAAATTTTGAATTCGAGTTT	3877				
Db	8264	GTTTTATGATTACAAAGTGATACACTAAATTTTATAAAATTTTGAATTCGAGTTT	8205			
3878	AATGGTGAATTCAGATCGAAAAAAGATGATGTTCTCTGACAAAGACGAGTAA	3937				
Db	8204	AATGGTGAATTCAGATCGAAAAAAGATGATGTTCTCTGACAAAGACGAGTAA	8145			
3938	AAAAATACAGATATGCGGAAACAAAGAGTTTTCAAAAATCTCGGTTGCGGCTTACG	3997				
Db	8144	AAAAATACAGATATGCGGAAACAAAGAGTTTTCAAAAATCTCGGTTGCGGCTTACG	8085			
3998	TATAGAAGATATCAAGAAAGGATCAGAACAAAAAATAAGCGAAAGCTCCGCTTT	4057				
Db	8084	TATAGAAGATATCAAGAAAGGATCAGAACAAAAAATAAGCGAAAGCTCCGCTTT	8025			
4058	TAGAGGATACAGATTTTCGCTACTGTTTGTATAGGTAA-TATATCATGGCTATTAA	4116				
Db	8024	TAGAGGATACAGATTTTCGCTACTGTTTGTATAGGTAA-TATATCATGGCTATTAA	7965			
4117	AAATACTAAGCTAGAAAAATTTGGATTTTATATATCTGACTCAATTCCTATGATTG	4176				
Db	7964	AAATACTAAGCTAGAAAAATTTGGATTTTATATATCTGACTCAATTCCTATGATTG	7905			
4177	GAAGAAAAATPAGAGATTTGGGCGTATCTATGGCTGTAGTCTTTTACAGTATGGA	4236				
Db	7904	GAAGAAAAATPAGAGATTTGGGCGTATCTATGGCTGTAGTCTTTTACAGTATGGA	7845			

QY	4237	CGAAAAAAGATATAAGATACATGAATAGTAGTGTGTTATACGAATCGAAAGCACTA	4296
Db	7844	CGAAAAAAGATATAAGATACATGAATAGTAGTGTGTTATACGAATCGAAAGCACTA	7785
QY	4297	TAAAAAACCCACATATACAGTTATATATGTCAGAAATCCTGTACAAATAGAAAGCGT	4356
Db	7784	TAAAAAACCCACATATACAGTTATATATGTCAGAAATCCTGTACAAATAGAAAGCGT	7725
QY	4357	TAGGAACAAGATTAAGCGAAAAATTTGGGAAATAGTTCAAGTTGCTCATGTTGAGTACTTGA	4416
Db	7724	TAGGAACAAGATTAAGCGAAAAATTTGGGAAATAGTTCAAGTTGCTCATGTTGAGTACTTGA	7665
QY	4417	TTATATCAAGGTTTCATATGAATATTTGACTCATGAATCAAAGGACGCTATGCTAGAA	4476
Db	7664	TTATATCAAGGTTTCATATGAATATTTGACTCATGAATCAAAGGACGCTATGCTAGAA	7605
QY	4477	TAAACATATATACGACAAAAAAGATATTTTGAAACATTAATGATTTGATTTGACCCGCTA	4536
Db	7604	TAAACATATATACGACAAAAAAGATATTTTGAAACATTAATGATTTGATTTGACCCGCTA	7545
QY	4537	TATAACACTTGTGTAAGCCAAAAAGAGAAATGAAGAATTTACTTTTGTAGATATAGTGA	4596
Db	7544	TATAACACTTGTGTAAGCCAAAAAGAGAAATGAAGAATTTACTTTTGTAGATATAGTGA	7485
QY	4597	TGACTATATTTTGTAAATACAAAGATTTTAATGCTTTTATTCGCTTAGCGGAGCGGA	4656
Db	7484	TGACTATATTTTGTAAATACAAAGATTTTAATGCTTTTATTCGCTTAGCGGAGCGGA	7425
QY	4657	GTTCGGAATTTTAAATACGAATGATGTAAGAATATTTGTTTCAACAACTCTAGCGCTT	4716
Db	7424	GTTCGGAATTTTAAATACGAATGATGTAAGAATATTTGTTTCAACAACTCTAGCGCTT	7365
QY	4717	TAGATTTAGTTTGGTGGGCAATATATCATGTGGATATAGCAAGTATGCAAGTTCT	4776
Db	7364	TAGATTTAGTTTGGTGGGCAATATATCATGTGGATATAGCAAGTATGCAAGTTCT	7305
QY	4777	TGATGCTGAAACGGGGGAAATTAATAATGACAAAGAAAGAGTATTTGCTGAAAAT	4836
Db	7304	TGATGCTGAAACGGGGGAAATTAATAATGACAAAGAAAGAGTATTTGCTGAAAAT	7245
QY	4837	GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTTGAAGATACAGAGAAATG	4896
Db	7244	GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTTGAAGATACAGAGAAATG	7185
QY	4897	GAGTTGAATTAAGTACACATAGATTTTATGAGGAGGAGGATTTGAATTAATAAAA	4956
Db	7184	GAGTTGAATTAAGTACACATAGATTTTATGAGGAGGAGGATTTGAATTAATAAAA	7125
QY	4957	GCCCCCTCGACAAAGTCCGCACTTCGTTCTTTTTT 4992	
Db	7124	G-CCCCCTCGACAAAGTCCGAGGGGTTTATTTTT 7090	

RESULT 12  
AY028776/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

AY028776 13561 bp DNA circular SYN 05-APR-2001  
TnphoZ mutagenesis vector pMHL120, complete plasmid sequence.  
AY028776  
AY028776.1 GI:13549398  
TnphoZ mutagenesis vector pMHL120  
TnphoZ mutagenesis vector pMHL120  
artificial sequences; vectors.  
1 (bases 1 to 13561)  
Lee M.H., Nittayajarn, A., Ross, R.P., Rothschild, C.B., Parsonage, D.,  
Claborn, A. and Rubens, C.E.  
Characterization of Enterococcus faecalis alkaline phosphatase and  
use in identifying Streptococcus agalactiae secreted proteins  
J. Bacteriol. 181 (18), 5790-5799 (1999)  
99412280  
10482522  
2 (bases 1 to 13561)  
Lee, M.H., Clancy, A., Jones, A.L., Nittayajarn, A. and Rubens, C.E.

TITLE Identification of genes encoding exported proteins in *Streptococcus agalactiae* using *tnphoZ* mutagenesis

## JOURNAL

Unpublished

3 (bases 1 to 13561)

## AUTHORS

Lee, M.H., Clancy, A., Jones, A.L., Nittayajarn, A. and Rubens, C.E.

## TITLE

Direct Submission  
Submitted (19-MAR-2001) Infectious Disease, Children's Hospital and Regional Medical Center, 4800 Sand Point Way NE, Seattle, WA 98105, USA

## FEATURES

## source

1. .13561  
/organism="tnphoZ mutagenesis vector pMHL120"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:156762"  
/lab\_host="Streptococcus agalactiae"  
/note="transposon *tnphoZ* delivery vehicle"  
complement(4167. .4928)  
/genes="erm"  
complement(4167. .4928)  
/genes="erm"  
/note="confers erythromycin-resistance"  
/codon\_start=1  
/transl\_table=11  
/product="rRNA methylase"  
/protein\_id="AAK27830.1"  
/db\_xref="GI:13549399"  
/translations="MYHFSDDVMNKIKYSONFLTNEKVLNOI IKOLNKETDVTVEI  
GTGKGLHTYKLAKISKQVSIYELDSHFNLSSEKLNIRVLIHQDILQFQPNKOR  
YKIVGNIPHLSTQIIRKVVFFSHASDIYLIVEEYFKRTLDIHTLGLLHTQVSIQ  
QLKLPACFCFQPKVNSVLIKLTHRTDVPDKYKLYTFVSKWNREYRQLETKNQ  
FHOAKHAKVNNLSVTYQVLSIFNSYLLFNGRK"  
complement(6372. .6563)  
/genes="rop"  
complement(6372. .6563)  
/genes="rop"  
/note="regulatory protein; involved in determination of  
plasmid copy number"

## gene

## CDS

1. .13561  
/transl\_table=11  
/product="ROP protein"  
/protein\_id="AAK27829.1"  
/db\_xref="GI:13549400"  
/translations="MTQKERTALNMARFIRSOITLLEKLNELDADEQADICESLHDH  
ADELYRSCLARFGDDGEN"  
complement(7201. .8391)  
/genes="tet"  
complement(7201. .8391)  
/genes="tet"  
/codon\_start=1  
/transl\_table=11  
/product="tetracycline-resistance protein"  
/protein\_id="AAK27830.1"  
/db\_xref="GI:13549401"  
/translations="MKSNKALIVLGTVTLDVAGLVMPVLPGLRDIIVHSDSTASH  
YGVLLALYALMFLCAPVLGALSDREGSPVLLASLGATIDYALMATPTVWLIVAG  
RIVAGITGATGAVAGYIADITDGEDRASHFGLMCAFCGVGAVGAGLLCAISLH  
APFLAAVLNGLLILQESHGKERRPMLRFPNVPVSRFWRGNTIIVAAUMTV  
FFTMQLGVQVPAALWIFGEDRPNSATMIGLSLVFGLHALAQAFVTPATKFEFG  
KQAI IAGRADALGYVLLAFATRGWMAFFIMILLASGGIGMPALQAMLSRQVDDHOG  
QLOGSLAALTSLTSITGPIVTAIYAASASTWNLAWIVGAALYVCLPALRRGAWSR  
ATST"

## gene

## CDS

1. .13561  
/transl\_table=11  
/product="alkaline phosphatase; derived from *Enterococcus faecalis*; expression is dependent upon integration event  
to create in-frame start codon"  
/allele="phoZ"  
10619. .11413  
/genes="aphA3"  
10619. .11413  
/genes="aphA3"  
/note="confers kanamycin-resistance"  
/codon\_start=1

## gene

## CDS

1. .13561  
/transl\_table=11  
/product="aminoglycoside phosphotransferase A3"  
/protein\_id="AAK27831.1"  
/db\_xref="GI:13549402"  
/translations="MAKMRISPELKKLIEKYRCVKDTGMSPAKVYKLVGENENLYLK  
MTDSRYKGTVDYVEREKDMLWLEGLPVKVLHFERHDSNLLMSEADGVLCSEY  
EDQSPKIIELVACICILPHSIDSDCPVNTSLDSELAELDYLLNLDADVDCEWE  
EDFPKDPRELYFLTEKEPEBELVSHGLDGSNIFVKDGKVGSGFIDJGRSGRADKW  
YDFAVCVRSIREDEEQYVELFVLGLGIRKPDWEKIKYIILLDELFF"  
complement(11817. .12515)  
/genes="repA3"  
complement(11817. .12515)  
/genes="repA3"  
/codon\_start=1  
/transl\_table=11  
/product="temperature-sensitive RepA"  
/protein\_id="AAK27832.1"  
/db\_xref="GI:13549403"  
/translations="MAIKNTKARNFGLYPDSIPNDWKEKLESIGYMAVSPHMD  
EKDKDTWNSNIQNGKHVKPHYVVIYARNPTVIESVRNKIKGLNNSVAVHVEI  
LDYIKSYEYLTHESDAIAKHHIYDKDILNINDEDDIVYITLDESOKREIKNILL  
DIVDDYNLNVKOLMFAFIRLGRGAEGILNVDKDIVSTNSSAFRLWFEGNYCGYRA  
SYAKVLDAETGEIK"

## gene

## CDS

1. .13561  
/transl\_table=11  
/product="aminoglycoside phosphotransferase A3"  
/protein\_id="AAK27831.1"  
/db\_xref="GI:13549402"  
/translations="MAKMRISPELKKLIEKYRCVKDTGMSPAKVYKLVGENENLYLK  
MTDSRYKGTVDYVEREKDMLWLEGLPVKVLHFERHDSNLLMSEADGVLCSEY  
EDQSPKIIELVACICILPHSIDSDCPVNTSLDSELAELDYLLNLDADVDCEWE  
EDFPKDPRELYFLTEKEPEBELVSHGLDGSNIFVKDGKVGSGFIDJGRSGRADKW  
YDFAVCVRSIREDEEQYVELFVLGLGIRKPDWEKIKYIILLDELFF"  
complement(11817. .12515)  
/genes="repA3"  
complement(11817. .12515)  
/genes="repA3"  
/codon\_start=1  
/transl\_table=11  
/product="temperature-sensitive RepA"  
/protein\_id="AAK27832.1"  
/db\_xref="GI:13549403"  
/translations="MAIKNTKARNFGLYPDSIPNDWKEKLESIGYMAVSPHMD  
EKDKDTWNSNIQNGKHVKPHYVVIYARNPTVIESVRNKIKGLNNSVAVHVEI  
LDYIKSYEYLTHESDAIAKHHIYDKDILNINDEDDIVYITLDESOKREIKNILL  
DIVDDYNLNVKOLMFAFIRLGRGAEGILNVDKDIVSTNSSAFRLWFEGNYCGYRA  
SYAKVLDAETGEIK"

## gene

## CDS

1. .13561  
/transl\_table=11  
/product="aminoglycoside phosphotransferase A3"  
/protein\_id="AAK27831.1"  
/db\_xref="GI:13549402"  
/translations="MAKMRISPELKKLIEKYRCVKDTGMSPAKVYKLVGENENLYLK  
MTDSRYKGTVDYVEREKDMLWLEGLPVKVLHFERHDSNLLMSEADGVLCSEY  
EDQSPKIIELVACICILPHSIDSDCPVNTSLDSELAELDYLLNLDADVDCEWE  
EDFPKDPRELYFLTEKEPEBELVSHGLDGSNIFVKDGKVGSGFIDJGRSGRADKW  
YDFAVCVRSIREDEEQYVELFVLGLGIRKPDWEKIKYIILLDELFF"  
complement(11817. .12515)  
/genes="repA3"  
complement(11817. .12515)  
/genes="repA3"  
/codon\_start=1  
/transl\_table=11  
/product="temperature-sensitive RepA"  
/protein\_id="AAK27832.1"  
/db\_xref="GI:13549403"  
/translations="MAIKNTKARNFGLYPDSIPNDWKEKLESIGYMAVSPHMD  
EKDKDTWNSNIQNGKHVKPHYVVIYARNPTVIESVRNKIKGLNNSVAVHVEI  
LDYIKSYEYLTHESDAIAKHHIYDKDILNINDEDDIVYITLDESOKREIKNILL  
DIVDDYNLNVKOLMFAFIRLGRGAEGILNVDKDIVSTNSSAFRLWFEGNYCGYRA  
SYAKVLDAETGEIK"

## ORIGIN

Query Match 31.8%; Score 1635.2; DB 12; Length 13561;  
Best Local Similarity 98.2%; Pred. No. 1.1e-262;  
Matches 1685; Conservative 0; Mismatches 28; Indels 3; Gaps 3;  
QY 3279 TTTATATTTTTCGGTATTTTATTAAAAAGCTCTCAAAATCGTTTCTGAGACGTTT 13285  
Db 13344 TATATATTTTATATCCGATTTTATTAAAAAGCTCTCAAAATCGTTTCTGAGACGTTT 13285  
QY 3339 TAGCGTTTATTCGTTTATCGGATATCGTAAACACGCGTTATCGTAGCGTAA 3398  
Db 13284 TAGCGTTTATTCGTTTATCGGATATCGTAAACACGCGTTATCGTAGCGTAA 13225  
QY 3399 AAGCCCTTGAGCGTAGCGT - GCTTTGAGCGAAGATGTTCTCTGTTAGATTATGAAGCC 3457  
Db 13224 AAGCCCTTGAGCGTAGCGTGGCTTTGAGCGAAGATGTTCTCTGTTAGATTATGAAGCC 13165  
QY 3458 GATGACATGAATGAATTAAGCGCGCGTCTTCTATTTTCGTTGAGGAGCGCTCAAGG 3517  
Db 13164 GATGACATGAATGAATTAAGCGCGCGTCTTCTATTTTCGTTGAGGAGCGCTCAAGG 13105  
QY 3518 GAGTTTGAGGGAATGAAATTCCTCATGGGTTTGAATTTAAAAATTCGTTCAATTTTCG 3577  
Db 13104 GAGTATGAGGGAATGAAATTCCTCATGGGTTTGAATTTAAAAATTCGTTCAATTTTCG 13045  
QY 3578 CGAGCGGTAGCGTGGAAAAATTTTGAAGAAATTTGGAATTTGGAAGAAATTTGCGGGGA 3637  
Db 13044 CGAGCGGTAGCGTGGAAAAATTTTGAAGAAATTTGGAATTTGGAAGAAATTTGCGGGGA 12985  
QY 3638 AAGGAAGCGAATTTTGCTTCGCTACTACGACCCCGCCCATTAAGTCCGAGTCCCAATTTT 3697  
Db 12984 AAGGAAGCGAATTTTGCTTCGCTACTACGACCCCGCCCATTAAGTCCGAGTCCCAATTTT 12925  
QY 3698 GTGCCAAAAACGCTCTATCCCACTGGCTCAAGGTTTGAGGGGTTTTCATTCGCCAAC 3757  
Db 12924 GTGCCAAAAACGCTCTATCCCACTGGCTCAAGGTTTGAAGGGGTTTTCATTCGCCAAC 12865  
QY 3758 GAATCGCAACGTTTTCGCCAACGTTTATATAATCTATATTAAGTACGTTTATTTGTT 3817  
Db 12864 GAATCGCAACGTTTTCGCCAACGTTTATATAATCTATATTAAGTACGTTTATTTGTT 12805  
QY 3818 GTTTTATGATTAACAAAGTGATACACPAATTTTATAAATTTTGAATTTGATTCGAGTTT 3877  
Db 12804 GTTTTATGATTAACAAAGTGATACACPAATTTTATAAATTTTGAATTTGATTCGAGTTT 12745  
QY 3878 AATCGTATTTTCAAGTTCGAAAAAAGAGTTATGATTTCTCTCACAAGAGCAAGATAA 3937  
Db 12744 AATCGTATTTTCAAGTTCGAAAAAAGAGTTATGATTTCTCTCACAAGAGCAAGATAA 12685



3938 AAAATTAACAGATATGCGGAAACAAAAAGTTTTCAAAATCTCGGTTGCGCGTTAGC 3997  
 12684 AAAATTAACAGATATGCGGAAACAAAAAGTTTTCAAAATCTCGGTTGCGCGTTAGC 12625  
 3998 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTT 4057  
 12624 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTT 12565  
 4058 TAGAAGGATACGAGTTTTCGCTACTTGTGTTTTCGATAGGTAA-TATATCATGCTATTAA 4116  
 12564 TAGAAGGATACGAGTTTTCGCTACTTGTGTTTTCGATAGGTAA-TATATCATGCTATTAA 12505  
 4117 AAATACTAAAGCTAGAAATTTTGGATTTTATATATCTGACTCAATTCCTAATGATTG 4176  
 12504 AAATACTAAAGCTAGAAATTTTGGATTTTATATATCTGACTCAATTCCTAATGATTG 12445  
 4177 GAAAGAAAATTAAGAGTTTGGCGTATCTATGCGTGCAGTCCCTTACAGATATGA 4236  
 12444 GAAAGAAAATTAAGAGTTTGGCGTATCTATGCGTGCAGTCCCTTACAGATATGA 12385  
 4237 CGAAAAAAGATAAGATACATGGAATAGTAGTGTATATACGAAATGGAAGCACTA 4296  
 12384 CGAAAAAAGATAAGATACATGGAATAGTAGTGTATATACGAAATGGAAGCACTA 12325  
 4297 TAAAAACACACTATACGTTATATATATGTCACGAAATCCTGTAACTAGAAAGCGT 4356  
 12324 TAAAAACACACTATACGTTATATATATGTCACGAAATCCTGTAACTAGAAAGCGT 12265  
 4357 TAGGAACAAGATTAAGCGAAATTTGGGAATAGTTTCAGTTGCTCATGTTGAGATACCTGA 4416  
 12264 TAGGAACAAGATTAAGCGAAATTTGGGAATAGTTTCAGTTGCTCATGTTGAGATACCTGA 12205  
 4417 TTATATCAAAAGGTTTCAATGAAATTTTGAATGATGATCAAGAGCGTATTCCTAAGAA 4476  
 12204 TTATATCAAAAGGTTTCAATGAAATTTTGAATGATGATGATCAAGAGCGTATTCCTAAGAA 12145  
 4477 TAAACATATATACGACAAAAAGATATTTGACATTAATGATTTGATTTACACGCTA 4536  
 12144 TAAACATATATACGACAAAAAGATATTTGACATTAATGATTTGATTTACACGCTA 12085  
 4537 TATAACACTTTCATGAAAGCCAAAAAGAGAAATGGAAGATTTACTTTTAGATATAGTGA 4596  
 12084 TATAACACTTTCATGAAAGCCAAAAAGAGAAATGGAAGATTTACTTTTAGATATAGTGA 12025  
 4597 TGACTATAATTTGTAATCAAAAGATTTAATGCGCTTTTATCGCCTTAGGGAGCGGA 4656  
 12024 TGACTATAATTTGTAATCAAAAGATTTAATGCGCTTTTATCGCCTTAGGGAGCGGA 11965  
 4657 GTTTGGAAATTTAAATACGAATGATGTAAGAGATTTGTTTCAACAACTCTAGCGCTT 4716  
 11964 GTTTGGAAATTTAAATACGAATGATGTAAGAGATTTGTTTCAACAACTCTAGCGCTT 11905  
 4717 TAGATTATGTTTGAGGCAATTTATCATGTTGGATATAGAGCAAGTTATGCAAGGTTCT 4776  
 11904 TAGATTATGTTTGAGGCAATTTATCATGTTGGATATAGAGCAAGTTATGCAAGGTTCT 11845  
 4777 TGATGCTTAACCGGGGAATTAATAATGCAACAAAGAAAGAGTTTATTGCTGAAAT 4836  
 11844 TGATGCTTAACCGGGGAATTAATAATGCAACAAAGAAAGAGTTTATTGCTGAAAT 11785  
 4837 GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAAATG 4896  
 11784 GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAAATG 11725  
 4897 GAAGTTGAATTAAGTACAACTAGATTTTATGAGAGGAGGATTTGATTAATAAAA 4956  
 11724 GAAGTTGAATTAAGTACAACTAGATTTTATGAGAGGAGGATTTGATTAATAAAA 11665  
 4957 GSCCCCTTGACGAAAGTCGCGACTTCGTTCTTTTTT 4992  
 11664 G-CCCCCTGACGAAAGTCGAGGGGTTTTTATTTT 11630

RESULT 13  
 LREPAB  
 LOCUS Lactococcus lactis plasmid replicon copx and repX genes. 2505 bp DNA linear BCT 20-MAY-1998  
 DEFINITION  
 ACCESSION X54310.1 GI:3150197  
 VERSION copx gene; replication protein; repX gene.  
 KEYWORDS Lactococcus lactis  
 SOURCE Lactococcus lactis  
 ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
 REFERENCE  
 1 Xu, F.F., Pearce, L.E. and Yu, P.L.  
 Genetic analysis of a lactococcal plasmid replicon  
 JOURNAL Mol. Gen. Genet. 227 (1), 33-39 (1991)  
 MEDLINE 91260676  
 PUBMED 1904536  
 2 Grohmann, E., Moscoso, M., Zechner, E.L., Espinosa, M. and del Solar, G.  
 In vivo definition of the functional origin of leading strand  
 replication of the Lactococcal plasmid pPX2  
 JOURNAL Unpublished  
 3 Yu, P.L.  
 Direct Submission  
 JOURNAL Submitted (08-AUG-1990) Yu P.-L., Massey University, Department of  
 Biotechnology, Palmerston North, New Zealand  
 4 (bases 1 to 2505)  
 DIRECT SUBMISSION  
 JOURNAL Submitted (18-MAY-1998) Espinosa M., Centro de Investigaciones  
 Biologicas, CSIC, Velazquez, 144. E-28006, Madrid, SPAIN  
 COMMENT On May 22, 1998 this sequence version replaced gi:44063.  
 FEATURES  
 Location/Qualifiers  
 1..2505  
 /organism="Lactococcus lactis"  
 /mol\_type="genomic DNA"  
 /strain="snp. lactis 5136"  
 /db\_xref="taxon:1358"  
 /plasmid="pPX2"  
 1054..1093  
 /rpt\_type=DIRECT  
 1087..1155  
 1211..1248  
 /rpt\_type=DIRECT  
 1289..1394  
 1312..1317  
 1329..1332  
 1341..1502  
 /gene="copX"  
 1341..1502  
 /gene="copX"  
 /codon\_start=1  
 /transl\_table=11  
 /product="COPX protein"  
 /protein\_id="CAA38207.1"  
 /db\_xref="GI:44064"  
 /db\_xref="GOA:O04136"  
 /db\_xref="SPTREMBL:O04136"  
 /translation="WVISEKRVMSLTKEQDKLTDWAKQKGFSAVAALAEY  
 AKSEEQKK"  
 1491..1552  
 /note="putative"  
 1498..1551  
 1555..1558  
 1569..1567  
 /gene="repX"  
 1569..1567  
 /gene="repX"  
 /codon\_start=1  
 /transl\_table=11  
 /product="RepX protein"  
 repeat\_region  
 rep origin  
 repeat\_region  
 -35\_signal  
 -10\_signal  
 RBS  
 gene  
 CDS



JOURNAL Plasmid 46 (2), 140-148 (2001)  
MEDLINE 21476064  
PUBMED 11591139  
REFERENCE 2 (bases 1 to 4438)  
AUTHORS Takamatsu, D., Osaki, M. and Sekizaki, T.  
TITLE Direct Submission  
JOURNAL Submitted (08-FEB-2001) Daisuke Takamatsu, National Institute of  
Animal Health, Laboratory of Molecular Bacteriology, Kannondai  
3-1-1, Tsukuba, Ibaraki 305-0856, Japan  
(E-mail: p1013d@nih.afrc.go.jp, Tel: 81-298-38-7743,  
Fax: 81-298-38-7743)  
FEATURES  
Location/Qualifiers  
1..4438  
/organism="Thermosensitive suicide vector pSET5s"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:152115"  
528..689  
/gene="orfC"  
528..689  
/gene="orfC"  
/codon\_start=1  
/transl\_table=11  
/product="OrfC"  
/protein\_id="BAB64884.1"  
/db\_xref="GI:15553036"  
/translation="MVISESKRWISLTKEQDKLIDMAKQKGFSAVALAIEY  
ARXSEOKK"  
756..1454  
/gene="repAts"  
756..1454  
/gene="repAts"  
/codon\_start=1  
/transl\_table=11  
/product="RepAts"  
/protein\_id="BAB64885.1"  
/db\_xref="GI:15553037"  
/translation="MAIKNTKARNFGLLYPDSIPNDWKEKLSLGVSMVSLPHMD  
EKQDWTNNISITQNGKHYKPHYIYIARNPVITSEVRNKKRLGSSVAHVEI  
LDYKSGVEYLTHESKDAIAKXKHIDYKDLINIDPDIDRYITLDESQKRELKLL  
DIVDYNLVNFKDMATIRLRGAEGFGLINTDVKDIVSTNSSAFRLWFEGNYCGYRA  
SYAKVLAETGEIK"  
complement(1803..2126)  
/gene="lacZ"  
complement(1803..2126)  
/gene="lacZ"  
/codon\_start=1  
/transl\_table=11  
/product="LacZ"  
/protein\_id="BAB64886.1"  
/db\_xref="GI:15553038"  
/translation="MTWITPFLHACRSTLSDPRVPSSNSLAVLQRRDWNPGVTQLN  
RLAAHPFPASWRNSEEARTRPSQQLSLNGEWMRYFLTLHLOGISHRIWCTLSTI  
CSDAA"  
3593..4243  
/gene="cat"  
3593..4243  
/gene="cat"  
/codon\_start=1  
/transl\_table=11  
/product="chloramphenicol acetyltransferase"  
/protein\_id="BAB64887.1"  
/db\_xref="GI:15553039"  
/translation="MNFNKIDLDWKEIFENHVLNQQTTFISITTEIDISVLNFKQ  
EGYKFPAPFIFLTVRINSNTAFRTGNSDGLGWDKPLFTIFDGVSKTFSGIWI  
PVKDFEFLDYLSDVEKYNGSKLPKTPENAFSLSIIPWTSFTGNLINNNS  
NYLLPIITAGKFINKNSIYLPFLSLQVHVSVDGYNHAGLFMNSIQELSDRPNWLL"  
ORIGIN  
Query Match 30.6%; Score 1575.2; DB 12; Length 4438;  
Best Local Similarity 98.7%; Pred. No. 1.3e-252;  
Matches 1619; Conservative 0; Mismatches 18; Indels 3; Gaps 3;  
QY 3355 TAGTTATCGGCATATCGTTAAACACAGCGGTATCGTAGCGTAAAGCCCTTGAGCGTAG 3414

3 TAGTTATCGGCATATCGTTAAACACAGCGGTATCGTAGCGTAAAGCCCTTGAGCGTAG 62  
3415 CCT-GCTTTGACGAGGAGATGTTCTGTAGATTATGAAGCCGATGACTGAATGAAT 3473  
63 CGTGGCTTTCAGCAGGAGATGTTCTGTAGATTATGAAGCCGATGACTGAATGAAT 122  
3474 AATAAGCGAGCGTCTCTATTTTCGGTTGGAGGAGGTCAAGGAGTTTGAGGGAATGA 3533  
123 AATAAGCGAGCGTCTCTATTTTCGGTTGGAGGAGGTCAAGGAGTTTGAGGGAATGA 182  
3534 AATTCCTCATCGGTTTGATTTTAAATAATGCTTCAATTTTGCAGCGGTAGCGCTGG 3593  
183 AATTCCTCATCGGTTTGATTTTAAATAATGCTTCAATTTTGCAGCGGTAGCGCTGG 242  
3594 AATAATTTTGAATAAATTTGAAATTTGAAAAAATGGGGGAAGGAAGCGAATTTTG 3653  
243 AATAATTTTGAATAAATTTGAAATTTGAAAAAATGGGGGAAGGAAGCGAATTTTG 302  
3654 CTTCCGTACTACGACCCGCCCATTAAGTGCAGAGTCCCAATTTTGTGCCAAAAACGCTCT 3713  
303 CTTCCGTACTACGACCCGCCCATTAAGTGCAGAGTCCCAATTTTGTGCCAAAAACGCTCT 362  
3714 ATCCCAACTGGCTCAAGGTTTGAGGGTTTTCATGCGCAACGAATCGCCAAAGCTTTT 3773  
363 ATCCCAACTGGCTCAAGGTTTAAAGGGTTTTCATGCGCAACGAATCGCCAAAGCTTTT 422  
3774 CGCCAAAGCTTTTATTAATCTATATTTAAAGTAGCTTTTATGTTGTTTATGATTACAA 3833  
423 CGCCAAAGCTTTTATTAATCTATATTTAAAGTAGCTTTTATGTTGTTTATGATTACAA 482  
3834 AGTGATACACTAATTTTATAAATTTATTTGATTTGAGTTTAAATGTTGATTTCAGAA 3893  
483 AGTGATACACTAATTTTATAAATTTATTTGATTTGAGTTTAAATGTTGATTTCAGAA 542  
3894 TCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAGATAAAAAATTAACAGATAG 3953  
543 TCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAGATAAAAAATTAACAGATAG 602  
3954 GCGAAAACAAAAAGTTTTTCAAAATCTCGGTTGCGGGTTAGCTATAGAAAGAAATGCA 4013  
603 GCGAAAACAAAAAGTTTTTCAAAATCTCGGTTGCGGGTTAGCTATAGAAAGAAATGCA 662  
4014 AGAAGGAATCAGACAAAAAATAAGCGAAGTCGCGTTTATAGAGGATAGAGTT 4073  
663 AGAAGGAATCAGACAAAAAATAAGCGAAGTCGCGTTTATAGAGGATAGAGTT 722  
4074 TTCGCTACTTTTGTATAAGGTAA-TATATCATCGTGGCTATTAAAAATACTAAAGCTAGA 4132  
723 TTCGCTACTTTTGTATAAGGTAA-TATATCATCGTGGCTATTAAAAATACTAAAGCTAGA 782  
4133 AATTTGGATTTTATATATCTGACTCAATCTCTAATGATTTGAAAGAAAAATTAGAG 4192  
783 AATTTGGATTTTATATATCTGACTCAATCTCTAATGATTTGAAAGAAAAATTAGAG 842  
4193 AGTTTGGGCTATCTATGCTCTGCTGCTTACAGATATGACGAAAAAAGATAAA 4252  
843 AGTTTGGGCTATCTATGCTCTGCTGCTTACAGATATGACGAAAAAAGATAAA 902  
4253 GATACATGGAATAGTAGTATGATATACGAAATGGAAGCACTATAAAAAACACACTAT 4312  
903 GATACATGGAATAGTAGTATGATATACGAAATGGAAGCACTATAAAAAACACACTAT 962  
4313 CAGCTTATATATTTGCGCAAAATCTGTAACAATAGAAAGCGTTAGGAACAAGATTAG 4372  
963 CAGCTTATATATTTGCGCAAAATCTGTAACAATAGAAAGCGTTAGGAACAAGATTAG 1022  
4373 CGAAAAATGGGAATAGTTTCAAGTTGCTCATGTTGAGATCTTGAATATCAAGGTTCA 4432  
1023 CGAAAAATGGGAATAGTTTCAAGTTGCTCATGTTGAGATCTTGAATATCAAGGTTCA 1082  
4433 TATGAATATTTGACTCATGAATCAAGGAGCGTTTCTTAAGNAATTAACATATACGAC 4492

```
Db 1083 TATGAATATTGACTCATGAATCAAGGACGCTATTGCTGAAGATAAACAATATACGAC 1142
Qy 4493 AAAAAAGATATTTTGAACATTAATGATTTTGAATGATTCACCGCTATATAACATGATGAA 4552
Db 1143 AAAAAAGATATTTTGAACATTAATGATTTTGAATGATTCACCGCTATATAACATGATGAA 1202
Qy 4553 AGCCAAAAGAGATTTGAAGATTTTACCTTTAGATATAGTACGATCATTAATTTGGTA 4612
Db 1203 AGCCAAAAGAGATTTGAAGATTTTACCTTTAGATATAGTACGATCATTAATTTGGTA 1262
Qy 4613 AATACAAAGATTTAATGCGCTTTTATCGCCTTAGGGAGCGGAGTTTGGAAATTTTAAAT 4672
Db 1263 AATACAAAGATTTAATGCGCTTTTATCGCCTTAGGGAGCGGAGTTTGGAAATTTTAAAT 1322
Qy 4673 ACGAATGATTAAGATATTTGTTTCAACAACTCTAGCGCCTTTAGATATAGTATGTTTGA 4732
Db 1323 ACGAATGATTAAGATATTTGTTTCAACAACTCTAGCGCCTTTAGATATAGTATGTTTGA 1382
Qy 4733 GGCAATTCAGTGTGATATAGCAAGTATGCAAGGTTCTTCTGATGCTGAACCGGG 4792
Db 1383 GGCAATTCAGTGTGATATAGCAAGTATGCAAGGTTCTTCTGATGCTGAACCGGG 1442
Qy 4793 GAAATAAATGACAAACAAAGAGAGTATTTGCTGAAATGAGGAATTTAAATAAG 4852
Db 1443 GAAATAAATGACAAACAAAGAGAGTATTTGCTGAAATGAGGAATTTAAATAAG 1502
Qy 4853 AAATTAGGACTTAAAGAGCGTATTCGAAAGATACAGAGAAATGGAATTAAGTA 4912
Db 1503 AAATTAGGACTTAAAGAGCGTATTCGAAAGATACAGAGAAATGGAATTAAGTA 1562
Qy 4913 CAACAATAGATTTATTCAGAGGAGGATTTTGAATAAATAAAGCCCTGACGAAG 4972
Db 1563 CAACAATAGATTTATTCAGAGGAGGATTTTGAATAAATAAAG-CCCCCTGACGAAG 1621
Qy 4973 TCGGACTTCGCTCTTTT 4992
Db 1622 TCGAAGGGGTTTATTTT 1641

RESULT 15
AB055652 4438 bp DNA circular SYN 21-MAY-2003
LOCUS Thermosensitive suicide vector pSET6s DNA, complete sequence.
ACCESSION AB055652
VERSION AB055652.1 GI:15553040
KEYWORDS Thermosensitive suicide vector pSET6s
SOURCE Thermosensitive suicide vector pSET6s
ORGANISM artificial sequences; vectors.
REFERENCE 1 Takamatsu, D., Osaki, M. and Sekizaki, T.
AUTHORS Thermosensitive suicide vectors for gene replacement in
TITLE Streptococcus suis
JOURNAL Streptococcus suis
MEDLINE Plasmid 46 (2), 140-148 (2001)
PUBMED 21476064
11591139
REFERENCE 2 (bases 1 to 4438)
AUTHORS Takamatsu, D., Osaki, M. and Sekizaki, T.
SUBMITTED Direct Submission
TITLE Submitted (08-FEB-2001) Daisuke Takamatsu, National Institute of
JOURNAL Animal Health, Laboratory of Molecular Bacteriology, Kannondai
3-1-1, Tsukuba, Ibaraki 305-0855, Japan
(E-mail: pl013@dnish.affrc.go.jp, Tel: 81-298-38-7743,
Fax: 81-298-38-7743)
FEATURES
Location/Qualifiers
1..4438
/organism="Thermosensitive suicide vector pSET6s"
/mol_type="genomic DNA"
/db_xref="taxon:152116"
528..689
/gene="orfC"
528..689
/gene="orfC"
969..1641
/gene="orfC"
gene
CDS
```

```

/codon_start=1
/transl_table=11
/product="orfC"
/protein_id="BAB64888.1"
/db_xref="GI:15553041"
/translation="MVISESKRVMISLTREQDKLTDMAKQKGFSAVALAIBEY
ARKESEOKK"
756..1454
/gene="repA"
756..1454
/gene="repA"
/codon_start=1
/transl_table=11
/product="repA"
/protein_id="BAB64889.1"
/db_xref="GI:15553042"
/translation="MAIKNTKARNFGFLYPDSIPNDWKEKLESLGVSMAVSLPHDMD
EKDKDTNNNSIIQNGKHYHYIYARNPVTIESVRNKIKRLGNSSVAHVVEI
LDYTKSYELTHESKDAIAKNKHIDKIDILNIDPDIDRYITLDESQREKLNLL
DIVDYNLVNTKDLMAFIRLGAEFGILNTNDVKDIVSTNSSAFRLWFEFGYCGYRA
SYAKVLAETGEIK"
complement (1803..2126)
/gene="lacZ"
complement (1803..2126)
/gene="lacZ"
/codon_start=1
/transl_table=11
/product="lacZ"
/protein_id="BAB64890.1"
/db_xref="GI:15553043"
/translation="MTWITSLHACRSTLEDPRVPSNSLAVLVLRDMDENPCVTOLN
RLAHPFPFASNRNSEARTDRPQQLRSLNGEWRLRYFLTLGLSHIRIWLSTFI
CSDAA"
complement (3572..4222)
/gene="cat"
complement (3572..4222)
/gene="cat"
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="BAB64891.1"
/db_xref="GI:15553044"
/translation="MNFNKIDLNWKKKEIFNHYLNQOTFTTTEIDISLVYRNKIQ
EGYKPYAFIPLATVINSNTAPRTGVNSDGLGYMDKLEPLATIPDGVSKTSGIWT
FVKDFEFIDLYLSDVEKNGSKLPKPTIPENAFSLIIPWTSFTGNLNNNS
NYLLPIITAGKFINKGNSTIYLPUSLQVHVSVDGYHAGLPMNSIQELSDRPDWLL"
Query Match 30.6%; Score 1575.2; DB 12; Length 4438;
Best Local Similarity 98.7%; Pred. No. 1.3e-252;
Matches 1619; Conservative 0; Mismatches 18; Indels 3; Gaps 3;
Qy 3355 TAGTTATCGGCATTAATCGTTAAACACGCGTTATCGTAGCGTAAAGCCCTTGAGCGTAG 3414
Db 3 TAGTTATCGGCATTAATCGTTAAACACGCGTTATCGTAGCGTAAAGCCCTTGAGCGTAG 62
Qy 3415 CGT-GCTTTGACGCGAAGATGTTGCTGTAGATTATGAAGCCGATGACTGAATGAAT 3473
Db 63 CTGCGCTTTGACGCGAAGATGTTGCTGTAGATTATGAAGCCGATGACTGAATGAAT 122
Qy 3474 AATAAGCCGACGCTCTCTATTTCGGTTGGAGGAGCTCAAGGAGTTTGAGGGAATGA 3533
Db 123 AATAAGCCGACGCGCTCTATTTCGGTTGGAGGAGCTCAAGGAGTTTGAGGGAATGA 182
Qy 3534 AATTCCTCATGGTTTGATTTTAAATTCCTTCAATTTTCCGAGCGGTAGCGTGG 3593
Db 183 AATTCCTCATGGTTTGATTTTAAATTCCTTCAATTTTCCGAGCGGTAGCGTGG 242
Qy 3594 AAAATTTTGAATAAATTTGAAATTTGAAAAAATTTGGGGGAAAGAGCGAATTTTG 3653
Db 243 AAAATTTTGAATAAATTTGAAATTTGAAAAAATTTGGGGGAAAGAGCGAATTTTG 302
Qy 3654 CTTCCGTACTAGACCCCTTAATGTCGCCGAGTGCCTATTTTGTGCAAAACGCTCT 3713
```

Db 303 CTTCGGTACTACGACCCCCCATTAAGTCCGAGTGCCCAATTTTGTGCGCAAAACGCTCT 362  
QY 3714 ATCCCACTGCTCTCAAGGGTTTGAGGGTTTTCAATCGCCACGAAATGCCCAACGTTTT 3773  
Db 363 ATCCCACTGCTCTCAAGGGTTTAAGGGTTTTCAATCGCCACGAAATGCCCAACGTTTT 422  
QY 3774 CGCCAAAGCTTTTATATAAATCTATATTAAAGTAGCTTTATTTGTTGTTTTATGATTACAA 3833  
Db 423 CGCCAAAGCTTTTATATAAATCTATATTAAAGTAGCTTTATTTGTTTTATGATTACAA 482  
QY 3834 AGTGATACACTAAATTTTATAAATTTTATGATGGAGTTTTTAAATGGTGATTCAGAA 3893  
Db 483 AGTGATACACTAAATTTTATAAATTTTATGATGGAGTTTTTAAATGGTGATTCAGAA 542  
QY 3894 TCGAAAAAAGAGTATGATTTCTCTGACAAAAGCAAGTAAAAAATTAACAGATATG 3953  
Db 543 TCGAAAAAAGAGTATGATTTCTCTGACAAAAGCAAGTAAAAAATTAACAGATATG 602  
QY 3954 GCGAAACAAAAGGTTTTTCAAATCTGCGGTTGCGCGTTAGCTATAGAAAGTAATGCA 4013  
Db 603 GCGAAACAAAAGGTTTTTCAAATCTGCGGTTGCGCGTTAGCTATAGAAAGTAATGCA 662  
QY 4014 AGAAGGAATCAGAACAAAATAAAGCGAAGCTCGGTTTTTGAAGGATACGAGTT 4073  
Db 663 AGAAGGAATCAGAACAAAATAAAGCGAAGCTCGGTTTTTGAAGGATACGAGTT 722  
QY 4074 TTGCTACTTGTGTTTGTATAGGTAA-TATATCATGCTATTAAAAATCTAAAGCTAGA 4132  
Db 723 TTGCTACTTGTGTTTGTATAGGTAAATATATCATGCTATTAAAAATCTAAAGCTAGA 782  
QY 4133 AATTTTGGATTTTATATATCTGACTCAATTCCTAATGATTGGAAGAAAAATTAGAG 4192  
Db 783 AATTTTGGATTTTATATATCTGACTCAATTCCTAATGATTGGAAGAAAAATTAGAG 842  
QY 4193 AGTTTGGCGTATCTATGGCTGTGAGTCCCTTTACCGATATGGACGAAAAAAGATAAA 4252  
Db 843 AGTTTGGCGTATCTATGGCTGTGAGTCCCTTTACCGATATGGACGAAAAAAGATAAA 902  
QY 4253 GATACATGGAATAGTAGTGTGTTATACGAAATGGAAGCACTATATAAAACCCACACTAT 4312  
Db 903 GATACATGGAATAATAGTAATATATATACAAATGGAAGCACTATATAAAACCCACACTAT 962  
QY 4313 CACGTTATATATGTCAGNAATCCCTGTAACATAGAAAGCGTTAGAACCAAGATTAG 4372  
Db 963 CACGTTATATATGTCAGNAATCCCTGTAACATAGAAAGCGTTAGAACCAAGATTAG 1022  
QY 4373 CGAAAAATGGCAATAGTTCAGTTGCTCATGTTGAGATCTTGATTATATCAAAGGTTCA 4432  
Db 1023 CGAAAAATGGCAATAGTTCAGTTGCTCATGTTGAGATCTTGATTATATCAAAGGTTCA 1082  
QY 4433 TATGAAATTTGACTCATGAATCAAAGGACGCTATTGCTAAGAAATAAACATATATACGAC 4492  
Db 1083 TATGAAATTTGACTCATGAATCAAAGGACGCTATTGCTAAGAAATAAACATATATACGAC 1142  
QY 4493 AAAAAAGATATTTGAACATTAATGATTGATATGACCGCTATATAACACTTTGATGAA 4552  
Db 1143 AAAAAAGATATTTGAACATTAATGATTGATATGACCGCTATATAACACTTTGATGAA 1202  
QY 4553 AGCCAAAAAGAGAAATGAAGAATTTACTTTTAGATATAGTGGATGACTATAATTTGGTA 4612  
Db 1203 AGCCAAAAAGAGAAATGAAGAATTTACTTTTAGATATAGTGGATGACTATAATTTGGTA 1262  
QY 4613 AATACAAAGATTTAATGGCTTTTATTCGCCCTTAGGGAGCGGAGTTTGGAAATTTAAAT 4672  
Db 1263 AATACAAAGATTTAATGGCTTTTATTCGCCCTTAGGGAGCGGAGTTTGGAAATTTAAAT 1322  
QY 4673 ACGAATGATGTAAGAATTTGTTTCAACAACTCTAGCGCTTTAGATTATGTTTGGAG 4732  
Db 1323 ACGAATGATGTAAGAATTTGTTTCAACAACTCTAGCGCTTTAGATTATGTTTGGAG 1382  
QY 4733 GCGAATTAATGATGGATATAGAGCAAGTTATGCAAGGTTCTTGATGCTGAAACCGGG 4792

Db 1383 GCGAATTTATCAGTGTGGATATAGACCAAGTTTATGCAAGGTTCTTGATGCTGAAACGGGG 1442  
QY 4793 GGAATATAATGACACAAACAAAGAAAGAGTTATTTGCTGAAATGAGGAATTAABAAAAG 4852  
Db 1443 GGAATATAATGACACAAACAAAGAAAGAGTTATTTGCTGAAATGAGGAATTAABAAAAG 1502  
QY 4853 AATTTAAGGACTTTAAAGAGCGTATTGAAAGATACAGAGAAATGGAAGTTGAAATTAAGTA 4912  
Db 1503 AATTTAAGGACTTTAAAGAGCGTATTGAAAGATACAGAGAAATGGAAGTTGAAATTAAGTA 1562  
QY 4913 CAACAATAGATTATTGAGAGGAGGATTTGAATAAATAAAGCCGCCCTGACGAAAG 4972  
Db 1563 CAACAATAGATTATTGAGAGGAGGATTTGAATAAATAAAG-CCCCCTGACGAAAG 1621  
QY 4973 TCGCGACTTCGTTCTTTTTT 4992  
Db 1622 TCGAAGGGGTTTTTTT 1641

Search completed: September 24, 2004, 04:19:41  
Job time : 12732.2 secs

This Page Blank (uspio)